

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: June 14, 2003, 19:19:30 ; Search time 10 seconds  
(without alignments)  
2836.981 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 635  
Sequence: 1 GGAGAGCTGAGGAGGATGTG.....CTTCAGTTAAATTATGTACA 342

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+.n2p.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool/US09518842/rnnat.14062003.175810.12259/app.query.fasta.1.519  
-DB=SwissProt.40 -OPMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -IOOPCL=0  
-IOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEASize=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09518842\_@CGN.1.1.26.@rnnat.14062003.175810.12259 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database: SwissProt.40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	619	97.5	139	1	INLA_HUMAN
2	141	22.2	182	1	RELX_PIG
3	116.5	18.3	185	1	RELX_MOUSE
4	111	17.5	178	1	RELH_RABIT
5	106.5	16.8	182	1	RELX_HORSE
6	105.5	16.6	166	1	REL2_PANTR
7	105.5	16.6	185	1	REL2_PANTR
8	103.5	16.3	185	1	REL2_HUMAN
9	103.5	16.3	185	1	REL1_HUMAN
10	100.5	15.8	185	1	RELX_MACMU
11	97.5	15.4	186	1	REL1_PANTR
12	76.5	12.0	160	1	RELX_RAT
13	75	11.8	160	1	RELX_CAVPO
14	75	11.8	836	1	DPOL_HPBPU
15	74.5	11.7	1531	1	PMPO_CHLTR
16	73.5	11.6	899	1	Y188_HUMAN
17	73	12.2	857	1	AR56_CANAL
18	73	12.2	993	1	RPNI_YEAST
			1395	1	SP41_YEAST

19	72.5	11.4	292	1	YHFR_SALTI	0824m8 salmonella
20	72.5	12.1	1744	1	CO4_HUMAN	P01028 homo sapien
21	72	11.3	3358	1	PGCV_MOUSE	062059 mus musculu
22	71.5	11.3	292	1	YFHR_SALTY	082n39 salmonella
23	70.5	11.1	271	1	VG18_BPP22	P03667 bacterioph
24	70.5	11.1	957	1	UVRA_BACSU	034863 bacillus su
25	70	10.9	283	1	PANC_SCHPO	009673 schizosacch
26	69.5	10.9	645	1	SVT_BACHD	09k866 bacillus ha
27	69	10.9	177	1	RELX_MESAU	064171 mesocricetu
28	69	11.5	274	1	RREP_VSVCM	P04878 vesicular s
29	69	11.5	943	1	BL14_CABEL	P51559 caenorhabdi
30	69	10.9	3396	1	PGCV_HUMAN	P13611 homo sapien
31	68	11.3	503	1	CP3P_MOUSE	009158 mus musculu
32	67.5	10.6	404	1	RPL1_TRYBG	P15593 trypanosoma
33	67	10.6	313	1	SDCL_RAT	P26260 rattus norv
34	67	11.2	449	1	YKE6_YEAST	P36091 saccharomyc
35	67	10.6	629	1	DXS_DEIRA	09rnb5 delinococcus
36	67	10.6	3828	1	TRX_DROVI	024742 drosophila
37	66.5	11.1	123	1	ELI2_PHYCR	P41803 phytopthor
38	66.5	10.5	295	1	YS86_MYCTU	010831 mycobacteri
39	66.5	10.5	979	1	T111_SCHPO	009813 schizosacch
40	66	11.0	897	1	CYRB_HUMAN	P32927 homo sapien
41	66	10.4	1025	1	CR2_MOUSE	P19070 mus musculu
42	66	11.0	3695	1	LMAS_HUMAN	015230 homo sapien
43	65.5	10.3	443	1	NRTA_SYNP7	P38043 synechococc
44	65.5	10.3	529	1	GAG_MLYDU	P23090 duplan murl
45	65.5	10.9	564	1	GAF2_SCHPO	010134 schizosacch

## ALIGNMENTS

RESULT 1  
INLA\_HUMAN  
ID INLA\_HUMAN STANDARD: PRT: 139 AA.  
AC Q14641.  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Early Placenta insulin-like peptide precursor (EPIL) (Placental)  
DE (Insulin-like peptide 4).  
GN INSL4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=96115599; PubMed=8666396;  
RA Chassin D., Laurent A., Jannet J.-L., Berger R., Bellet D.;  
RT "Cloning of a new member of the insulin gene superfamily (INSL4)  
RT expressed in human placenta.";  
RL Genomics 29:465-470(1995).  
RN [2]  
RP PROCESSING.  
RX MEDLINE=97430657; PubMed=9284764;  
RA Bellet D., Lavalssiere L., Mock P., Laurent A., Sabourin J.C.,  
RA Begosa P., Le Bouellier P., Frydman R., Troalen F., Bidart J.M.;  
RT "Identification of pro-EPIL and EPIL peptides translated from  
RT insulin-like 4 (INSL4) mRNA in human placenta.";  
RL J. Clin. Endocrinol. Metab. 82:3169-3172(1997).  
RN [3]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=98411035; PubMed=9740319;  
RA Laurent A., Rouillac C., Delezoid A.L., Giovannardi Y., Vekemans M.,  
RA Bellet D., Abitbol M., Vidaud M.;  
RT "Insulin-like 4 (INSL4) gene expression in human embryonic and  
RT trophoblastic tissues.";  
RL Mol. Reprod. Dev. 51:123-129(1998).  
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN TROPHOBLAST DEVELOPMENT  
CC AND IN THE REGULATION OF BONE FORMATION.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA, UTERUS AND IN FETAL  
CC PERICHONDRIUM.



CC or send an email to license@isb-sib.ch).

DR EMBL: K01088; AAA31114.1; -

DR EMBL: J02792; AAA31115.1; -

DR EMBL: A16593; CAA01295.1; -

DR EMBL: A06852; CAA00600.1; -

DR PIR: A01615; RXPG

DR PIR: A29796; A29796.

DR PDB: 1RLX; 15-OCT-94.

DR PDB: 2RLX; 15-OCT-94.

DR PDB: 3RLX; 15-OCT-94.

DR PDB: 4RLX; 15-OCT-94.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin; 1.

DR SMART: SM00078; IIGF; 1.

DR PROSITE: PS00262; INSULIN; 1.

KW Insulin family; Hormone; Signal; 3D-structure.

FT SIGNAL 1 24

FT CHAIN 25 56

FT PROPEP 57 154

FT CHAIN 161 182

FT MOD\_RES 25 25

FT DISULFID 34 169

FT DISULFID 46 182

FT DISULFID 168 173

FT CONFLICT 47 47

FT CONFLICT 51 44

FT CONFLICT 116 170

FT CONFLICT 170 170

SO SEQUENCE 182 AA; 20818 MW; 20736EB089F13AB4 CRC64;

Alignment Scores:

Score: 1.42e-07 Length: 182

Percent Similarity: 141.00 Matches: 41

Best Local Similarity: 40.00% Conservative: 7

Query Match: 22.20% Mismatches: 32

Indels: 40

Gaps: 2

US-09-518-842-1\_COPY\_76\_417 (1-342) x RELX\_PIG (1-182)

QY 100 CTGGAATCTGACGTCCTCAAGAAATGCTGTCACCTCCACACACAAAGATGACAGCC 159

DB 63 LeuGIuThrGlyProProIaIaGluThrMetProSerSerIleThrLysAspIaIaGluIle 82

QY 160 TTAGGTAGACATTCAGAAATTCATTCCTAATTTGTCCAGCAGAGCTGAGAAACACTGTCT 219

DB 83 LeuLysMetMetLeuGluThrValProAsnLeuProGlnGluLeuLysAlaThrLeuSer 102

QY 220 GAAGGGCAGCCATTCATTCG----- 237

DB 103 GluArgGlnProSerLeuArgGluLeuGlnGlnSerAlaSerLysAspSerAsnLeuAsn 122

QY 238 -----AGAAAAATATACTT----- 252

DB 123 PheGluGluPheLysLysIleIleLeuAsnArgGlnAsnGluAlaGluAspLysSerLeu 142

QY 253 -----TCCGCCCAAAAAGAGAAGTGCAGCTCAC 279

DB 143 LeuGluLeuLysAsnLeuGlyLeuAspLysHisSerArgLysLysArgLeuPheArgMet 162

QY 280 AGATTTCATTCCTGTTGTTGGAAGTAATTTGTGACATGACATTCCTGATTAAATTTATCT 339

DB 163 ThrLeuSerGluLysCysCysGlnValGlyLysIleArgLysAspIleAlaArgLeuLys 182

RESULT 3

RELX\_MOUSE

ID RELX\_MOUSE STANDARD; PRT; 185 AA.

AC P47932;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Prorelaxin precursor.

GN RUN OR RLX.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Swiss; TISSUE=Ovary;

RX MEDLINE=93196663; PubMed=8452637;

RA Evans B.A., John M., Fowler K.J., Summers R.J., Cronk M.,

RA Shine J., Tregear G.W.;

RT "The mouse relaxin gene: nucleotide sequence and expression.";

RT J. Mol. Endocrinol. 10:15-23(1993).

CC -!- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN

CC TO PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS.

CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Z27088; CAA81611.1; -

DR MGD: MGI:97931; R1N.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin; 1.

DR SMART: SM00078; IIGF; 1.

DR PROSITE: PS00262; INSULIN; FALSE\_NEG.

KW Insulin family; Hormone; Signal.

FT SIGNAL 1 22

FT CHAIN 23 57

FT PROPEP 58 156

FT CHAIN 161 185

FT DISULFID 36 171

FT DISULFID 48 185

FT DISULFID 170 175

SO SEQUENCE 185 AA; 20570 MW; 2B9E753EB8B5087B CRC64;

Alignment Scores:

Score: 7.13e-05 Length: 185

Percent Similarity: 116.50 Matches: 38

Best Local Similarity: 35.76% Conservative: 16

Query Match: 18.35% Mismatches: 56

Indels: 41

Gaps: 2

US-09-518-842-1\_COPY\_76\_417 (1-342) x RELX\_MOUSE (1-185)

QY 7 CTGAGGGGATGTGCTCCCGCATTTGGAAACACTTGCTGCTCATATTGCCCATGCTGAG 66

DB 33 IleArgMetCysGlyArgGluTyrAlaArgGluLeuLysIleCysGlyAlaSerVal 52

QY 67 AAGACATTCACACACACACCCAGAGAGGTGGCTGCTGGAATCTGAGCTCCCAAGAATA 126

DB 53 GlyArgLeuAlaLeuSerGlnGluProAlaLeuLeuAlaArgGlnAlaThrGluVal 72

QY 127 GTGTCACCTCCCAACAAAGATGGACAGACCTTGTAGTACGACATTCATTCATTCCT 186

DB 73 ValProSerPheIleAsnLysAspAlaGluProPheAspThrThrLeuLysCysLeuPro 92

QY 187 AATTGTCACACAGCTGAAGAAACCAACCATGCTGGAAGGCGACCATTCATTCGAAGAAATA 246

DB 93 AsnLeuSerGluLeuLysAlaValLeuSerGluAlaGlnAlaSerLeuProGluLeu 112

QY 247 -----ATACTT----- 252

DB 113 GlnHisAlaProValLeuSerAspSerValSerLeuGluGlyPheLysLysThrLeu 132

QY 252 ----- 252

Db 133 HisAspArgLeuclYgluaIaAspGlySerProProGlyLeuLysTyrLeuGlnSer 152

QY 253 -----TCCGGCAAAAGAGAGACTGGACGTCACAGATTGGTCATTCGTTGTGA 303

Db 153 AspThrHisSerTrgylsYsArgGluSerGlyGlyLeuMetSerGlnGlnCysCysHis 172

QY 304 GAAATTTGTGACGATGACATTCAGTTAAATTA 336

Db 173 ValGlyCysSerArgArgSerIleAlaLysLeu 183

RESULT 4

RELH\_RABIT STANDARD; PRT; 178 AA.

ID RELH\_RABIT STANDARD; PRT; 178 AA.

AC P51456;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Relaxin-like protein S010 precursor.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RC TISSUE=Tracheobronchial epithelium;

RX MEDLINE=93002619; PubMed=1339318;

RA Jettan A.M., Bernacki S.H., Floyd E.E., Saunders N.A., Pleniazek J.,

RA Jettan R.;

RT "Expression of a preprorelaxin-like gene during squamous

RT differentiation of rabbit tracheobronchial epithelial cells and its

RT suppression by retinoic acid."

RT Cell Growth Differ. 3:549-556(1992).

RL -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Secreted (BY SIMILARITY).

CC -1- INDUCTION: DURING SQUAMOUS CELL DIFFERENTIATION. REPRESSED BY

CC RETINOIC ACID.

CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: S45940; AMB23648.1; "

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin. 1.

DR SMART: SM00078; IIGF. 1.

KW PROSITE: PS00262; INSULIN. 1.

KW Insulin family; Hormone; Signal.

FT SIGNAL 1 20

FT CHAIN 21 52

FT PROPEP 53 150

FT CHAIN 155 178

FT CHAIN 155 178

FT DISULFID 34 165

FT DISULFID 46 178

FT DISULFID 164 169

SO SEQUENCE 178 AA; 20294 MW; F6A54D98A6B53211 CRC64;

Alignment Scores:

Pred. No.: 0.00286 Length: 178

Score: 111.00 Matches: 39

Percent Similarity: 32.81% Conservative: 3

Best Local Similarity: 30.47% Mismatches: 42

Query Match: 17.48% Indels: 44

DB: 1 Gaps: 3

US-09-518-842-1\_COPY\_76\_417 (1-342) x RELH\_RABIT (1-178)

QY 37 CACTTGTGTCTCATTTATTCCTCCATGCTGAGAAAGACATTCACACCCAGAGGGTGG 96

Db 50 HisLeuGluArgGluSerProSerProGluAsnProPhe----- 62

QY 97 CTGCTGGAATGTGGACGTCCCAAGAAATGTGTCACACCTCCCAACAAAGATGGACAA 156

Db 63 --LeuSerSerGlyProAlaAlaGlnThrValProSerSerIleLysLysAspAlaGln 81

QY 157 GCCTTAGGTACGACATCATTCATTCCTAATTTGTCCACAGAGCTGAGAAACCACTG 216

Db 82 AsnAlaAsnThrMetLeuGluSerIleProAsnLeuProGlnGluLeuThrAlaThrLeu 101

QY 217 TCTGAGAGGCGACCATCA----- 234

Db 102 PheGluLysGlnProSerLysLeuTyrLeuGlnTyrLeuProThrLeuLysSerAsn 121

QY 235 -----TTGAGAAATAATATA----- 249

Db 122 ValSerPheGluGluPheLysLysIleIleGlnAsnIleGlnArgGlyValGlnGlySer 141

QY 250 -----CTTCCCGCAAAAGAGAGAGTGGACGTACACGATTTTGTAT 288

Db 142 SerAlaSerGluSerAsnThrPheSerArgLysLysArgGlnPheSerGluSerLeuPro 161

QY 288 CCATTCTGTGTGAGAGTATTTGT 312

Db 162 GluGluCysCysLysTyrGlyCys 169

RESULT 5

RELX\_HORSE STANDARD; PRT; 182 AA.

ID RELX\_HORSE STANDARD; PRT; 182 AA.

AC P22969; Q28907;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Prerelaxin precursor (RXN).

GN RLN.

OS Equus caballus (horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI\_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hokkaido; TISSUE=Placenta;

RA Min K., Shiota K., Ogawa T.;

RT "Molecular cloning of equine preprorelaxin cDNA."

RT J. Reprod. Dev. 42:171-178(1996).

RN [2]

RP SEQUENCE OF 32-174 FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=95359320; PubMed=7543295;

RA Klonisch T., Ryan P.L., Yamashiro S., Porter D.G.;

RT "Partial complementary deoxyribonucleic acid cloning of equine relaxin

RT messenger ribonucleic acid, and its localization within the equine

RT placenta."

RT Biol. Reprod. 52:1307-1315(1995).

RN [3]

RP SEQUENCE OF 26-53 AND 163-182.

RC TISSUE=Placenta;

RX MEDLINE=91275796; PubMed=2055195;

RA Stewart D.R., Nevins B., Hadas E., Vandlen R.;

RT "Affinity purification and sequence determination of equine relaxin."

RT Endocrinology 129:375-383(1991).

CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN

CC TO PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS

CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000201; BAA19069.1; -
DR EMBL: S78800; AAB35036.1; -
DR PIR: B49739; B49739.
DR PIR: B49739; B49739.
DR HSSP: P01348; IRLX.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF000049; Insulin; 1.
DR SMART: SM00078; IGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
DR Insulin family; hormone; signal.
FT SIGNAL 1 25
FT CHAIN 26 53
FT PROPEP 54 156
FT CHAIN 161 182
FT DISULFID 35 169
FT DISULFID 47 182
FT DISULFID 168 173
FT DISULFID 173 168
FT CONFLICT 66 66
FT CONFLICT 133 133
SQ SEQUENCE 182 AA; 20721 MW; E5C9414303A838B8 CRC64;

Alignment Scores:
Pred. No.: 0.000901 Length: 182
Score: 106.50 Matches: 30
Percent Similarity: 52.94% Conservative: 15
Best Local Similarity: 35.29% Mismatches: 33
Query Match: 16.77% Indels: 7
DB: 1 Gaps: 3

US-09-518-842-1_COPY_76_417 (1-342) x RELX_HORSE (1-182)
QY 7 CTGAGGGAGTGTGTCCTCCGATTTGGAACACCTTGATTCGATTCGCT--- 63
DB 32 ILEYSALACYSGLIYARGLULEUALAARGLEUALGILLGILLCYSGILSERLEUSER 51
QY 64 ---GAGAGACATTC-----ACACACACCCGAGGAGGCTGCTGCTGAATCTGGAGCT 114
DB 52 TTPLYSLYSTHVALLEUARGLLEUGLUGLUGLUGLUGLUGLUGLUGLUGLUGL 68
QY 115 CCCAAGAAATGCTGTCACCTCCCAACAAACAAAGATGACACAGCTTAGTACGACATCA 174
DB 69 PROVALGILUILEVALSERISERISERISERISERISERISERISERISERISERISER 88
QY 175 GAATTCATTCTCAATTTGTCACACGAGCTGGAACAAACACCTGCTGAAGGCGACCATCA 234
DB 89 GIYLEAUSNSERASLEUPROLYSGLUINLUSALATHTLSEUSERGLUARGINPROSER 108
QY 235 TTGAGAGAAATATA 249
DB 109 TTPARGLULEUL 113

RESULT 6
REL2_PANTR
ID REL2_PANTR STANDARD: PRT: 166 AA.
AC PS1455; P79267;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prorelaxin H2 precursor (Fragment).
GN RNL2 OR RNLX2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;

```

```

RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISSUE=Placenta;
RX MEDLINE=94238260; PubMed=8182365;
RA Evans B.A., Fu P., Tregear G.W.;
RT "Characterization of two relaxin genes in the chimpanzee.";
RL J. Endocrinol. 140:385-392(1994).
RN [2]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2), AND TISSUE SPECIFICITY.
RC TISSUE=Placenta, and ovary;
RX MEDLINE=96328899; PubMed=8735594;
RA Gunneren J.M., Fu P., Roche P.J., Tregear G.W.;
RT "Expression of human relaxin genes: characterization of a novel
RT alternatively-spliced human relaxin mRNA species.";
RL Mol. Cell. Endocrinol. 118:85-94(1996).
CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
CC PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
CC INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY.
CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CORPUS LUTEUM OF PREGNANCY
CC AND IN THE PLACENTA.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z27245; CAAB1758.1; -
DR EMBL: S83209; AAD14430.1; -
DR HSSP: P04090; 6RLX.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF000049; Insulin; 1.
DR SMART: SM00078; IGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
DR Insulin family; hormone; Multigene family; signal; polymorphism;
KW Alternative splicing.
FT NON_TER 1 1
FT SIGNAL <1 5
FT CHAIN 6 34
FT PROPEP 37 138
FT CHAIN 143 166
FT DISULFID 16 153
FT DISULFID 28 166
FT DISULFID 152 157
FT VARSPLIC 52 84
FT FT
FT VARSPLIC 85 166
FT FT
FT VARSPLIC 30 34
FT VARIANT 30 34
SQ SEQUENCE 166 AA; 18760 MW; 220851E3134CEBDB CRC64;

Alignment Scores:
Pred. No.: 0.00114 Length: 166
Score: 105.50 Matches: 43
Percent Similarity: 35.67% Conservative: 13
Best Local Similarity: 27.39% Mismatches: 44
Query Match: 16.61% Indels: 57
DB: 1 Gaps: 6

US-09-518-842-1_COPY_76_417 (1-342) x REL2_PANTR (1-166)
QY 1 GCAAGAGCTGAGGGAGTGTGCT---CCCGCATTTGGAACACCTTGCTGCATATTTGCC 57
DB 23 ALACINLEALALECYSGILYSERTHTTIPSERLYSARGSERLEUSERGINLUASP 42

```

QY	58	ATCCGGAAGACATTCACCCACCCAGAGGGTGGCTGGAATCGAGCCGCC	117
Db	43	AlaProGln-----ThrPro-----ArgPro	49
QY	118	---AAGAAATGGTGTCACCTCCACAACAAGAATGGACAGCCTTAGTCAGCATCA	174
Db	50	ValAlaGluIleValAlaProSerPheIleAsnIleGlnThrIleAsnMetMetSer	69
QY	175	GAATTCATTCCTTAATTGTCACCAAGCTGGAAGAACCACTGTCTGAAGGGCACCACATCA	230
Db	70	GluPheValAlaAsnLeuProGlnLeuLysLeuThrLeuSerGluMetGlnProAla	89
QY	235	TTG-----	237
Db	90	LeuProGlnLeuGlnGlnIleThrValProValLeuLysAspSerSerLeuLeuPheGlnGlu	109
QY	238	---AAGAAATAATAACTT-----	252
Db	110	PheLysLysLeuIleArgAsnArgInserGlnAlaIleAspSerSerProSerGluLeu	128
QY	253	-----TCCGCCAAAAGAGACAGTGAACGTCACACAGATTGGAT	288
Db	130	LysThrLeuGlnLeuAspThrHisSerArgLysLysArgGlnLeuThrSerAlaLeuAla	149
QY	289	CCATTCTGGTGGAACTAATTCTGTGACGATGGAACCTCAGTTAATTAATTATGT	339
Db	150	AsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys	166

## RESULT 7

ID	REL2_HUMAN	STANDARD;	PRT;	185 AA.
AC	P04090; Q9UCX3; Q99936;			
DT	01-NOV-1986 (Rel. 03)	Created		
DT	01-NOV-1986 (Rel. 03)	Last sequence update		
DT	15-JUN-2002 (Rel. 41)	Last annotation update		
DE	Prorelaxin H2 precursor.			
GN	RLN2.			
OS	Homo sapiens (Human).			
OC	Euryarchaeota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RX	SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.			
RX	MEDLINE=85051298; PubMed=6548702;			
RA	Gudson P., John M., Crawford R., Haralambidis J., Scanlon D.,			
RA	Hudson J., Tregear G., Shine J., Niall H.;			
RT	"Relaxin gene expression in human ovaries and the predicted structure			
RL	of a human preprorelaxin by analysis of cDNA clones.";			
RL	EMBO J. 3:2353-2359(1984).			
RN	[2]			
RX	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RA	Serra H.;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	PARTIAL SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE			
RP	SPECIFICITY.			
RC	TISSUE=prostate;			
RX	MEDLINE=96328899; PubMed=8735594;			
RA	Gumersen J.M., Fu P., Roche P.J., Tregear G.M.;			
RT	"Expression of human relaxin genes: characterization of a novel			
RL	alternatively-spliced human relaxin mRNA species.";			
RL	Mol. Cell. Endocrinol. 118:85-94(1996).			
RN	[4]			
RP	SYNTHESIS.			
RX	MEDLINE=91250367; PubMed=2040595;			
RA	Buellesbach E.E., Schwabe C.;			
RT	"Total synthesis of human relaxin and human relaxin derivatives by			
RL	solid-phase peptide synthesis and site-directed chain combination.";			
RL	J. Biol. Chem. 266:10754-10761(1991).			
RN	[5]			
RP	PARTIAL SEQUENCE OF 25-51.			
RC	TISSUE=Semen;			

RA	MEDLINE=92241162; Pubmed=1572287.
RA	Winslow J.W., Shih A., Bourrell J.H., Weiss G., Reed B., Stults J.T.,
RA	Goldsmith L.T.,
RT	"Human seminal relaxin is a product of the same gene as human luteal
RT	relaxin."
RL	Endocrinology 130:2660-2668(1992).
RP	[6]
RP	SEQUENCE OF 25-53 AND 162-185.
RX	MEDLINE=91167739; Pubmed=2076664;
RA	Winslow J.W., Griffin P.R., Rinderknecht E., Vandlen R.L.;
RT	"Structural characterization by mass spectrometry of native and
RT	recombinant human relaxin."
RL	Biomed. Environ. Mass Spectrom. 19:655-664(1990).
RN	[7]
RP	X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX	MEDLINE=92015205; Pubmed=1656049;
RA	Eigenbrodt C., Randal M., Quan C., Burnier J., O'Connell L.,
RA	Rinderknecht E., Kossiakoff A.A.;
RT	"X-ray structure of human relaxin at 1.5 A. Comparison to insulin and
RT	implications for receptor binding determinants."
RL	J. Mol. Biol. 221:15-21(1991).
CC	-1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
CC	PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
CC	INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY,
CC	PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
CC	-1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC	SULFIDE BONDS.
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC	produced by alternative splicing.
CC	-1- TISSUE SPECIFICITY: Expressed in the ovary during pregnancy. Also
CC	expressed in placenta, decidua and prostate.
CC	-1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL: X00948; CAA25460.1; -
DR	EMBL: AL135786; CAC04177.1; -
DR	EMBL: AL135786; CAC04176.1; -
DR	EMBL: S83200; AAD14429.1; -
DR	EMBL: A17315; CAA01324.1; -
DR	EMBL: A06925; CAA00602.1; -
DR	PIR: A60982; A60982.
DR	PDB: 6RLX; 31-OCT-93.
DR	Genev: HGNC:10027; RLN2.
DR	MIM: 179740; -
DR	InterPro: IPR004825; Ins/IGF/relax.
DR	Pfam: PF00049; Insulin; 1.
DR	SMART: SM00078; IIGF; 1.
DR	PROSITE: PS00262; INSULIN; 1.
KW	Insulin family; Hormone; Multigene family; Signal; 3D-structure;
KW	Alternative splicing.
FT	SIGNAL 1 24
FT	CHAIN 25 53
FT	PROPEP 56 157
FT	CHAIN 162 185
FT	MOD_RES 162 162
FT	DISULFID 35 172
FT	DISULFID 47 185
FT	DISULFID 171 176
FT	VARSPLOC 71 117
FT	VARSPLOC 117 117
FT	VARSPLOC 118 185
FT	HELIx 27 29
FT	STRAND 31 32
FT	HELIx 36 49
FT	HELIx 36 49



FT VARSPLIC 71 117 EIVPSTINKDTEITITIMELFIANLPKELKAALSEROPSIDE  
 FT LQOYVP -> GDEIQTIVSLGISPDGKALRTSCFTREPLG  
 FT ALSKLYPSSTKIOKL (IN ISOFORM 2).  
 FT VARSPLIC 118 185 MISSING (IN ISOFORM 2).  
 FT VARIANT 28 28 K -> M (IN DBSNP:618066).  
 FT SEQUENCE 185 AA; 21145 MW; B318628ABFE07142 CRC64;

## Alignment Scores:

Pred. No.: 0.00193 Length: 185  
 Score: 103.50 Matches: 44  
 Percent Similarity: 35.67% Conservative: 12  
 Best Local Similarity: 28.03% Mismatches: 44  
 Query Match: 16.30% Indels: 57  
 Gaps: 6

US-09-518-842-1\_COPY\_76\_417 (1-342) x RELX\_HUMAN (1-185)

QY 1 GCAGAGCTGAGGGGATGTGCT---CCCCGATTGGGAACACTTGTCTCATATTGCCCC 57  
 DB 42 AlaglnlilealalecysglymetserthrtrpserlysrgrserleuSerGlnGluasp 61  
 QY 58 ATGCGTGAAGAGACATTACACACACCCAGAGGGTGGCTGCTGGAATCTGAGCGTCCC 117  
 DB 62 Alaprolin-----ThrPro-----ArgPro 68  
 QY 118 ---AAGAAATGCTGTACACCTCCACACAAAGATGACACCTTAGTACGACATCA 174  
 DB 69 ValalaglilevalproserPheileasnllyasphrthrlellellellelle 88  
 QY 175 GAATTCATCTCTATTGTGTACACAGACCTGAGAAACCACTGCTGTGAGGCGACCATCA 234  
 DB 89 GlupheilealaleasnlleuProProGlnleuLysAlaAlaLeuSerGlnGlnProser 108  
 QY 235 TTG----- 237  
 DB 109 LeuSerGlnleuGlnGlnInTyValProAlaLeuLysasPserAsnLeuSerPheGlnGlu 128  
 QY 238 ---AAGAAATTAATCTTCCCGC----- 258  
 DB 129 PheLysLysLeuIleArgAsnArgGlnSerGlnAlaAlaasPserAsnProSerGlnleu 148  
 QY 259 -----AAAAAGAGAGTGGAGCTGACAGATTGGAT 288  
 DB 149 LysTyrleuGlyLeuAspThrHisSerGlnLysLysArgProGlyValAlaLeuPhe 168  
 QY 289 CCATTCTGTGTGGAAGTAATTGTGTGACAGTGAACCTTCAATTAATTATGT 339  
 DB 169 GluLysCysCysLeuIleGlyCysThrLysArgSerleuAlaLysTyrCys 185

## RESULT 9

RELX\_MACMU STANDARD; PRT: 185 AA.

AC P19884;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Proteolaxin precursor.  
 GN RLN.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90073957; PubMed=2590381;  
 RA Crawford R.J., Hammond V.E., Roche P.J., Johnston P.D., Tregear G.W.;  
 RT "Structure of rhesus monkey relaxin predicted by analysis of the  
 single-copy rhesus monkey relaxin gene."  
 RL J. Mol. Endocrinol. 3:169-174(1989).  
 CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO  
 PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE

CC INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY,  
 CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.  
 CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO  
 CC DISULFIDE BONDS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
 CC PIR: A34936; A34936.  
 DR HSP, P04090; GRX.  
 DR InterPro: IPR004825; Ins/IGF/relax.  
 DR Pfam: PF00049; Insulin; 1.  
 DR SMART: SM00078; IIGF; 1.  
 DR PROSITE: PS00262; INSULIN; 1.  
 KW Insulin family; Hormone; Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 53 RELAXIN B CHAIN (PROBABLE).  
 FT PROPEP 56 157 CONNECTING PEPTIDE (PROBABLE).  
 FT CHAIN 161 185 RELAXIN A CHAIN (PROBABLE).  
 FT DISULFID 35 172 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 47 185 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 171 176 BY SIMILARITY.

SEQ SEQUENCE 185 AA; 20895 MW; 7E3C5D21B57E185C CRC64;

## Alignment Scores:

Pred. No.: 0.00193 Length: 185  
 Score: 103.50 Matches: 41  
 Percent Similarity: 34.39% Conservative: 13  
 Best Local Similarity: 26.11% Mismatches: 46  
 Query Match: 16.30% Indels: 57  
 Gaps: 5

US-09-518-842-1\_COPY\_76\_417 (1-342) x RELX\_MACMU (1-185)

QY 1 GCAGAGCTGAGGGGATGTGCT---CCCCGATTGGGAACACTTGTCTCATATTGCCCC 57  
 DB 42 AlaglnlilealalecysglymetserthrtrpserlysrgrserleuSerGlnGlnasp 61  
 QY 58 ATGCGTGAAGAGACATTACACACACCCAGAGGGTGGCTGCTGGAATCTGAGCGTCCC 117  
 DB 62 AlaprolleuLysPro-----ArgPro 68  
 QY 118 ---AAGAAATGCTGTACACCTCCACACAAAGATGACACCTTAGTACGACATCA 174  
 DB 69 AlalaglilevalproserleuLysnGlnasphrthrlellellellellelle 88  
 QY 175 GAATTCATCTCTATTGTGTACACAGCTGGAAGAAACCACTGCTGAGGCGACCATCA 234  
 DB 89 GluphevalalaleasnlleuProGlnleuLysLysThrLeuSerGlnArgGlnProala 108  
 QY 235 TTG----- 237  
 DB 109 LeuSerGlnleuGlnGlnInHisValProValleuLysasPserAsnLeuSerPheGlnGlu 128  
 QY 238 ---AAGAAATTAATCTT----- 252  
 DB 129 PheLysLysIleleArgLysArgGlnSerGlnAlaThrAspSerSerProSerGlnleu 148  
 QY 253 -----TCCGCAAAAAGAGAGTGAACCTTCAATTAATTATGT 339  
 DB 149 ArgSerleuGlyLeuAspThrHisSerArgLysArgGlnLeuTyrMetThrLeuSer 168  
 QY 289 CCATTCTGTGTGGAAGTAATTGTGTGACAGTGAACCTTCAATTAATTATGT 339  
 DB 169 AsnLysCysGlnHisIleGlyCysThrLysLysSerleuAlaLysPheCys 185

## RESULT 10

RELX\_PANTR STANDARD; PRT: 166 AA.

AC P51454;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Proteolaxin H1 precursor (Fragment).  
 GN RNLI OR RLXI.



OS Pan troglodytes (Chimpanzee).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=94238260; PubMed=8182365;  
 RA Evans B.A., Fu P., Tregear G.W.;  
 RT "Characterization of two relaxin genes in the chimpanzee.";  
 RL J. Endocrinol. 140:385-392(1994).  
 CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO  
 CC PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE  
 CC INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY,  
 CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.  
 CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO  
 CC DISULFIDE BONDS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CORPUS LUTEUM OF PREGNANCY  
 CC BUT NOT IN THE PLACENTA.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 227225; CAA81739.1; -  
 DR HSSP: P04090; 6RLX.  
 DR InterPro: IPR004825; Ins/IGF/relax.  
 DR Pfam: PF00049; Insulin; 1.  
 DR SMART: SM00078; IIGF; 1.  
 DR PROSITE: PS00262; INSULIN; 1.  
 KW Insulin family; Hormone; Multigene family; Signal.  
 FT FT NON TER 1 1  
 FT STGNAL <1 5  
 FT CHAIN 6 34  
 FT PROPEP 37 139  
 FT CHAIN 143 166  
 FT DISULFID 16 153  
 FT DISULFID 28 166  
 FT DISULFID 152 157  
 SQ SEQUENCE 166 AA; 18730 MW; 7F469B1FB9259F4F CRC64;  
 Alignment Scores:  
 Pred. No.: 0.00407 Length: 166  
 Score: 100.50 Matches: 41  
 Percent Similarity: 36.948 Conservative: 17  
 Best Local Similarity: 26.118 Mismatches: 42  
 Query Match: 15.83% Indels: 57  
 DB: 1 Gaps: 7  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x REL1\_PANTR (1-166)  
 QY 1 GCAGAGCTGAGGGGAGTGTGTC---CCCGATTGGAAAAACACTGCTGTCATTATTCGCC 57  
 DB 23 AAlAGInleleallleCysglYmetserlhrtrpserlYsArGserleuSerglInluasp 42  
 QY 58 ATGCCGTGAGAAACATTCACCCACCCAGAGGGTGGCTGCTGGAATTCGAGAGTCCC 117  
 DB 43 AlAPrOgIn-----ThrPro-----ArgPro 49  
 QY 118 ---AAAGAAATGCTCAACCTCCACAAACAAGATGACAGCTTAGTACGACATCA 174  
 DB 50 ValAlaGluIleValAlProSerPheIleAsnYlsAspThrGluThrIleIleIleMetleu 69  
 QY 175 GAATCATTCCTCAATTTGTGCACACAGCTGAAGAACAACACTGTCTGAAGGCGACCATCA 234  
 DB 70 GluPheIleAlaAsnleuProGluIleuYlsAlaIleuSergluArgGlnProSer 89

QY 235 TMG----- 237  
 DB 90 LeuProGluProGlnGlnIleTyValProAlaIleuYlsAspSerAsnleuSerglGlu 109  
 QY 238 ---AAGAAATTAATCTTCCCGCAAAAAGAGAGT----- 270  
 DB 110 PheYlsYlsleuIleArgAsnArgGlnSergluAlaIleAspSerAsnProSerglGlu 129  
 QY 271 -----GGAGCTCACAGATTTCATTCATTC----- 294  
 DB 130 LysTYrleuGlyleuAspThrHisSerglInYlsYsArgGlnProTyValAlaIleuPhe 149  
 QY 295 -----TGTGTGAGAGTAATTTGTGCACGATGACGACTTCACTTAATTAATG 339  
 DB 150 GluYsCysCysleuIleGlyCysThrYlsArgSerleuAlaAsnTYrCys 166  
 RESULT 11  
 RELX\_RAT  
 ID RELX\_RAT STANDARD; PRT; 186 AA.  
 AC P01347;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, last sequence update)  
 DE 16-OCT-2001 (Rel. 40, last annotation update)  
 GN Proteolaxin precursor.  
 GN RLN.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81197624; PubMed=7231533;  
 RA Hudson P., Haley J., Cronk M., Shine J., Niall H.;  
 RT "Molecular cloning and characterization of cDNA sequences coding for  
 RT rat relaxin.";  
 RL Nature 291:127-131(1981).  
 RN [2]  
 RP SEQUENCE OF 23-57 AND 163-188.  
 RX MEDLINE=81090283; PubMed=7004862;  
 RA John M.J., Borjesson B.W., Walsh J.R., Niall H.D.;  
 RT "Limited sequence homology between porcine and rat relaxins:  
 RT implications for physiological studies.";  
 RL Endocrinology 108:726-729(1981).  
 CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN  
 CC TO PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS.  
 CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO  
 CC DISULFIDE BONDS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J00780; AAA24029.1; -  
 DR EMBL: V01264; CAA24578.1; -  
 DR PIR: A01614; RXRT.  
 DR HSSP: P01348; 1RLX.  
 DR InterPro: IPR004825; Ins/IGF/relax.  
 DR Pfam: PF00049; Insulin; 1.  
 DR SMART: SM00078; IIGF; 1.  
 DR PROSITE: PS00262; INSULIN; 1.  
 KW Insulin family; Hormone; Signal.  
 FT FT STGNAL 1 22  
 FT CHAIN 23 57  
 FT PROPEP 58 158  
 FT CHAIN 163 186  
 FT MOD\_RES 163 163  
 FT DISULFID 36 173  
 RELAXIN B CHAIN.  
 CONNECTING PEPTIDE.  
 RELAXIN A CHAIN.  
 PYROLIDONE CARBOXYLIC ACID.  
 INTERCHAIN (BY SIMILARITY).



OC Viruses; Retroviral viruses; Hepadnaviridae; Avihepadnavirus.  
 OX NCBI\_TaxID=12639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84138772; PubMed=669938;  
 RA Mandart E., Kay A., Gilbert F.;  
 RT "Nucleotide sequence of a cloned duck hepatitis B virus genome:  
 RT comparison with woodchuck and human hepatitis B virus sequences.";  
 RN J. Virol. 49:782-792(1984).  
 [2]  
 RP SEQUENCE OF 794-836 FROM N.A.  
 RX MEDLINE=84216498; PubMed=6328037;  
 RA Molnar-Kimber K.L., Summers J.W., Mason W.S.;  
 RT "Mapping of the cohesive overlap of duck hepatitis B virus DNA and of  
 RT the site of initiation of reverse transcription.";  
 RN J. Virol. 51:181-191(1984).  
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 CC + [DNA](N).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: K01834; AAA5742.1; -  
 DR PIR: A00710; JNVDL.  
 DR InterPro: IPR001463; DNAPOL\_viral\_C.  
 DR InterPro: IPR000201; DNAPOL\_viral\_N.  
 DR InterPro: IPR000477; RVISE.  
 DR Pfam: PF00078; rvc\_1.  
 DR Pfam: PF00242; DNA\_pol\_viral\_N\_1.  
 DR Pfam: PF00336; DNA\_pol\_viral\_C\_1.  
 DR ProDom: PD000814; DNAPOL\_viral\_C\_1.  
 KW .transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;  
 KW Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.  
 SQ SEQUENCE 836 AA; 95275 MW; FFEBS7CFE5494F4 CRC64;  
 Alignment Scores:  
 Pred. No.: 3 46 Length: 836  
 Score: 75.00 Matches: 25  
 Percent Similarity: 42.42% Conservative: 17  
 Best Local Similarity: 25.25% Mismatches: 27  
 Query Match: 11.81% Indels: 30  
 DB: 1 Gaps: 6  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x DPOL\_HPBVD (1-836)  
 OY 14 GATGTGTCCTCCGATTTGGAACACCTTCTCATATTCGCCATGCTGAGA----- 67  
 DB 295 AspmetValArgLInValSerAsnThrArgSerArgValArgAlaAsnGly 314  
 OY 68 -----AGCATTCACACACCCAGAGAGGT-----GGTGGTGGAACTGGAGCTC 115  
 DB 315 GlyAspLysHis-----ProProGlnSerGlySerLeuAlaCysTrp--GlyGlyLysG 332  
 OY 116 CCAAGAAATGGTGTCAACCTCCACACCAAGATGAGCAACACCTTAGTACGACATCAG 175  
 DB 332 InSerArgLLeuLysSerAspSerArgAspSerSerAla----- 346  
 OY 176 AATTCATTCCTAATTTGTACACGAGCTGAGAAACCACTG--TCTGAGGCGACGCAT 232  
 DB 347 -----ProValAspSerArgGlyArgProL 355  
 OY 233 CATGAGAAATATA-----CTTCCCGCAAAAGAGAGAGCTGAC 279  
 DB 355 ySerThrArgSerPheSerProLeuSerArgArgLysThrThrGlyAsnHis 372  
 RESULT 14

PMPD\_CHLTR  
 ID PMPD\_CHLTR STANDARD: PRT: 1531 AA.  
 AC 084818;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable outer membrane protein pmpd precursor (Polymorphic membrane  
 DE protein D).  
 GN PMPD OR CTR12.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatunov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)  
 CC (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001353; AAC68408.1; -  
 DR PHCT-2DPAGE; 084818; -  
 DR InterPro: IPR003368; Chlamydia\_PMP.  
 DR InterPro: IPR003357; OMP.  
 DR Pfam: PF02385; OMP\_1.  
 DR Pfam: PF02415; DUF145; 1.  
 KW Outer membrane; Signal; Multigene family; Complete proteome.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 1531 PROBABLE OUTER MEMBRANE PROTEIN PMPD.  
 SQ SEQUENCE 1531 AA; 160747 MW; 584ABD11E241EF22 CRC64;  
 Alignment Scores:  
 Pred. No.: 4 35 Length: 1531  
 Score: 74.50 Matches: 29  
 Percent Similarity: 44.79% Conservative: 14  
 Best Local Similarity: 30.21% Mismatches: 26  
 Query Match: 11.73% Indels: 27  
 DB: 1 Gaps: 5  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x PMPD\_CHLTR (1-1531)  
 OY 4 GAGCTGAGGGGATGTGTCCTCCGATTTGGAACACCTTG----- 42  
 DB 472 GlnThrGlnGlyGlyAlaLeuPheGlyGlnAlaSerLeuSerGluAsnAlaGly 491  
 OY 43 -----CTGTCATATTCCTCCCATGCTGAGAAACATTCACCCACACC-----CCA 87  
 DB 492 ValLeuThrPheLysAspAsnIleValLysThrPheAlaSerAsnGlyLysIleLeuGly 511  
 OY 88 GGAGGTGGCTGCGAATCTGAGATCGACCCCAAGAAATGCTGCAACCTCCACCAACAA 147  
 DB 512 GlyLysAlaIleLeuAlaThrGlyLys-----ValGluIleThrAsnAsnSer 527  
 OY 148 GATGGA-----CAAGCTTAGTACGACATCAGAA 177  
 DB 528 GluGlyLysSerPheThrGlyAsnAlaArgAlaProGlnAlaLeuProThrGlnGlu 547  
 OY 178 TTCATTCCTAATTTGTACACGAGCTGAGAAACCACTGCTCAAGG 225



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: June 14, 2003, 19:43:16 ; Search time 12.5 Seconds  
(without alignments)  
1610.022 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417  
Perfect score: 635  
Sequence: 1 GCAGAGCTGACGAGGATGTCG.....CTTCAGTTAAATATGATGACA 342

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODSL=frame\_plus.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/US09518842/runat\_14062003\_175812.12308/app.query.fasta\_1.519  
-DB=Issued Patents\_AA -QEMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -STRAP=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09518842 -ECGN\_1.1.28.@runat\_14062003\_175812.12308 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCtUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	619	97.5	124	US-08-991-890-5	Sequence 15, Appl1
2	619	97.5	139	US-08-950-720A-15	Sequence 15, Appl1
3	619	97.5	139	US-08-991-890-2	Sequence 2, Appl1
4	619	97.5	139	US-08-991-890-4	Sequence 4, Appl1
5	613	96.5	139	US-09-599-564A-2	Sequence 2, Appl1
6	608	95.7	174	US-09-174-465D-2	Sequence 2, Appl1
7	255	40.2	51	US-09-174-465D-10	Sequence 10, Appl1
8	255	40.2	51	US-09-599-564A-10	Sequence 10, Appl1
9	193	30.4	41	US-09-174-465D-8	Sequence 8, Appl1
10	193	30.4	41	US-09-599-564A-8	Sequence 8, Appl1
11	193	30.4	41	US-09-201-227A-27	Sequence 27, Appl1
12	166	26.1	30	US-09-174-465D-12	Sequence 12, Appl1

13	166	26.1	30	US-09-599-564A-12	Sequence 12, Appl1
14	147	23.1	25	US-09-201-227A-28	Sequence 28, Appl1
15	105.5	16.6	150	US-08-443-568B-12	Sequence 12, Appl1
16	105.5	16.6	150	PCT-US94-06997-12	Sequence 12, Appl1
17	105.5	16.6	162	5464756-18	Patent No. 5464756
18	105.5	16.6	185	US-08-950-720A-12	Sequence 12, Appl1
19	103.5	16.3	185	US-08-950-720A-13	Sequence 13, Appl1
20	99.5	15.7	164	5464756-20	Patent No. 5464756
21	73	12.2	994	US-08-699-103B-3	Sequence 3, Appl1
22	73	12.2	994	US-08-229-059-3	Sequence 3, Appl1
23	71.5	11.3	778	US-09-556-877-193	Sequence 193, Appl1
24	71.5	11.3	778	US-09-620-412C-193	Sequence 193, Appl1
25	71.5	11.3	1530	US-09-556-877-178	Sequence 178, Appl1
26	71.5	11.3	1530	US-08-620-412C-178	Sequence 178, Appl1
27	69	10.9	2409	5180808-2	Patent No. 5180808
28	67	10.6	313	US-08-078-683A-4	Sequence 4, Appl1
29	66.5	10.5	846	US-09-134-001C-3067	Sequence 3067, Appl1
30	66	11.0	433	US-09-046-158A-2	Sequence 2, Appl1
31	66	11.0	897	US-07-960-389-2	Sequence 2, Appl1
32	65.5	10.3	1183	US-09-134-001C-3530	Sequence 3530, Appl1
33	64.5	10.2	156	US-08-188-582-18	Sequence 5102, Appl1
34	64.5	10.2	704	US-09-134-001C-5102	Sequence 18, Appl1
35	64.5	10.2	704	US-08-646-715-18	Sequence 18, Appl1
36	64	10.7	188	US-09-332-934-2	Sequence 2, Appl1
37	63.5	10.0	383	US-09-134-001C-3068	Sequence 3068, Appl1
38	63.5	10.0	572	US-08-669-524-8	Sequence 8, Appl1
39	63.5	10.0	619	US-09-423-890-4	Sequence 4, Appl1
40	63	9.9	446	US-08-244-205-5	Sequence 5, Appl1
41	63	9.9	446	US-09-161-994A-5	Sequence 5, Appl1
42	63	9.9	446	PCT-US92-10284-5	Sequence 5, Appl1
43	63	9.9	446	PCT-US94-01321-10	Sequence 10, Appl1
44	63	9.9	535	US-09-269-731-4	Sequence 4, Appl1
45	62.5	9.8	554	US-08-524-051-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-991-890-5  
Sequence 5, Application US/08991890  
Patent No. 6114307  
GENERAL INFORMATION:  
APPLICANT: Jaspers, Stephen R.  
APPLICANT: Sprugel, Katherine H.  
APPLICANT: Ren, Hong Ping  
APPLICANT: Humes, Jacqueline M.  
APPLICANT: Hoffman, Ross C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zymogenetics, Inc.  
STREET: 1201 Eastlake Avenue East  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,890  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/033,003  
FILING DATE: December 16, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sawisak, Deborah A  
REGISTRATION NUMBER: 37,438

REFERENCE/DOCKET NUMBER: 96-41  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-442-6672  
TELEFAX: 206-442-6678  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 124 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-991-890-5

Alignment Scores:  
Pred. No.: 3,43e-69 Length: 124  
Score: 619.00 Matches: 114  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.48% Indels: 0  
DB: 3 Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-08-991-890-5 (1-124)

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTGGAAACACTTGCTCATATTGCCCATG 60  
|||  
DB 11 AlaGlueuArGlyCysGlyProArGPhGlyShLsLeuSerTyCysPromet 30  
QY 61 CCGAGAAGACATTCACACACCCAGAGGGTGGCTGCTGAATCTGGACGTCCCAA 120  
|||  
DB 31 ProGluYsThrPheThrThrProGlyGlyTyrPleuLeuGluSerGlyArGProLys 50  
QY 121 GAATGGTGTCACTCCCAACAAGATGGACCAAGCTTAGGTACAGATCAGATTC 180  
|||  
DB 51 GlueValSerThrSerAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 70  
QY 181 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTGAAG 240  
|||  
DB 71 IleProAsnLeuSerProGluLeuLysProLeuSerGluGlnProSerLeuLys 90  
QY 241 AAATATATCTTCCCGCAAAAAGAGAGTGCACAGATTGATTCATTCCTGTTGT 300  
|||  
DB 91 LysIleIleLeuSerArGlyLysArgSerGlyArGHisArgPheAspPropheCys 110  
QY 301 GAAGTAATTTGTGACAGATGGAATCTCAGTTAATTATGTACA 342  
|||  
DB 111 GluValIleCysAspAspGlyThrSerValLysLeuGlyThr 124

RESULT 2  
US-08-950-720A-15  
Sequence 15, Application US/08950720A  
Patent No. 6046028  
GENERAL INFORMATION:  
APPLICANT: Conklin, Darrell C.  
APPLICANT: Lofton-Day, Catherine E.  
APPLICANT: Lok, Si  
APPLICANT: Jaspers, Stephen R.  
TITLE OF INVENTION: INSULIN HOMOLOG  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zymogenetics, Inc.  
STREET: 1201 Eastlake Avenue East  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/950.720A  
FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sawislak, Deborah A.  
REGISTRATION NUMBER: 37,438  
REFERENCE/DOCKET NUMBER: 96-09  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-442-6672  
TELEFAX: 206-442-6678  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6046028e  
US-08-950-720A-15

Alignment Scores:  
Pred. No.: 3,59e-69 Length: 139  
Score: 619.00 Matches: 114  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.48% Indels: 0  
DB: 3 Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-08-950-720A-15 (1-139)

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTGGAAACACTTGCTCATATTGCCCATG 60  
|||  
DB 26 AlaGlueuArGlyCysGlyProArGPhGlyShLsLeuSerTyCysPromet 45  
QY 61 CCGAGAAGACATTCACACACCCAGAGGGTGGCTGCTGAATCTGGACGTCCCAA 120  
|||  
DB 46 ProGluYsThrPheThrThrProGlyGlyTyrPleuLeuGluSerGlyArGProLys 65  
QY 121 GAATGGTGTCACTCCCAACAAGATGGACCAAGCTTAGGTACAGATCAGATTC 180  
|||  
DB 66 GlueValSerThrSerAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85  
QY 181 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTGAAG 240  
|||  
DB 86 IleProAsnLeuSerProGluLeuLysProLeuSerGluGlnProSerLeuLys 105  
QY 241 AAATATATCTTCCCGCAAAAAGAGAGTGCACAGATTGATTCATTCCTGTTGT 300  
|||  
DB 106 LysIleIleLeuSerArGlyLysArgSerGlyArGHisArgPheAspPropheCys 125  
QY 301 GAAGTAATTTGTGACAGATGGAATCTCAGTTAATTATGTACA 342  
|||  
DB 126 GluValIleCysAspAspGlyThrSerValLysLeuGlyThr 139

RESULT 3  
US-08-991-890-2  
Sequence 2, Application US/08991890  
Patent No. 6114307  
GENERAL INFORMATION:  
APPLICANT: Jaspers, Stephen R.  
APPLICANT: Sprugel, Katherine H.  
APPLICANT: Ren, Hong Ping  
APPLICANT: Humes, Jacqueline M.  
APPLICANT: Hoffman, Ross C.  
APPLICANT: Conklin, Darrell C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zymogenetics, Inc.  
STREET: 1201 Eastlake Avenue East  
CITY: Seattle

STATE: WA  
 COUNTRY: USA  
 ZIP: 98102  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/991,890  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/033,003  
 FILING DATE: December 16, 1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sawislak, Deborah A  
 REGISTRATION NUMBER: 37,438  
 REFERENCE/DOCKET NUMBER: 96-41  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-442-6672  
 TELEFAX: 206-442-6678  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 139 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FRAGMENT TYPE: internal  
 US-08-991-890-2

Alignment Scores:  
 Pred. No.: 3,596-69  
 Score: 619.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 97.488  
 DB: 3  
 Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-08-991-890-2 (1-139)

QY 1 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAACACTTCTCATATTGCCCCATG 60  
 DB 26 AAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeuSerTyrCysPromet 45  
 QY 61 CCTGAGAACACATTGACACACACCCAGAGGGGTGGCTGAGATTCGTGAGCGTCCCAA 120  
 DB 46 ProGluLysThrPheThrThrThrProGlyGlyTrpLeuLeuGluSerGlyArgProLys 65  
 QY 121 GAAATGGTGTCACTCCACACAAAGATGACAAAGCCTTAGTCGACATCAGAAATTC 180  
 DB 66 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85  
 QY 181 ATTCTTAATTTGTCACGAGCTGAGAGAACCACTGTCTGAAGGCGACCATCATTTGAG 240  
 DB 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyGlnProSerLeuLys 105  
 QY 241 AAAATAATACTTTCCTCCGAAAAAGAGAGTGGACGTCAAGATTGATTCATTCGTGTGT 300  
 DB 106 LysIleIleLeuSerArgLysArgSerGlyArgHisArgPheAspPropheCysCys 125  
 QY 301 GAAGTAATTTGTGACGATGAGCACTTCAGTTAATTATGATAC 342  
 DB 126 GluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

RESULT 4  
 US-08-991-890-4  
 Sequence 4, Application US/08991890  
 Patient No. 6114307  
 GENERAL INFORMATION:  
 APPLICANT: Jaspers, Stephen R.

APPLICANT: Sprugel, Katherine H.  
 APPLICANT: Ren, Hong Ping  
 APPLICANT: Humes, Jacqueline M.  
 APPLICANT: Hoffman, Ross C.  
 APPLICANT: Conklin, Darrell C.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: ZymoGenetics, Inc.  
 STREET: 1201 Eastlake Avenue East  
 CITY: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98102

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/991,890  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/033,003  
 FILING DATE: December 16, 1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sawislak, Deborah A  
 REGISTRATION NUMBER: 37,438  
 REFERENCE/DOCKET NUMBER: 96-41  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-442-6672  
 TELEFAX: 206-442-6678  
 TELEX:

INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 159 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FRAGMENT TYPE: internal  
 US-08-991-890-4

Alignment Scores:  
 Pred. No.: 3,796-69  
 Score: 619.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 97.488  
 DB: 3  
 Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-08-991-890-4 (1-159)

QY 1 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAACACTTCTGTCAATATGCCCCATG 60  
 DB 46 AAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeuSerTyrCysPromet 65  
 QY 61 CCTGAGAACACATTGACACACACCCAGAGGGGTGGCTGAGATTCGTGAGCGTCCCAA 120  
 DB 66 ProGluLysThrPheThrThrThrProGlyGlyTrpLeuLeuGluSerGlyArgProLys 85  
 QY 121 GAAATGGTGTCACTCCACACAAAGATGACAAAGCCTTAGTCGACATCAGAAATTC 180  
 DB 86 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 105  
 QY 181 ATTCTTAATTTGTGACGAGCTGAGAGAACCACTGTCTGAAGGCGACCATCATTTGAG 240  
 DB 106 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyGlnProSerLeuLys 125  
 QY 241 AAAATAATACTTTCCTCCGAAAAAGAGAGTGGACGTCAAGATTGATTCATTCGTGTGT 300

Db 126 LysIleIleLeuSerArgIlySarSerGIyArgHISArgPheAspProPheCysCys 145  
QY 301 GAAGTAATTGTGACGATGACCTGACGTAATATGATGACA 342  
Db 146 GluValIleIleCysAspAspGlyThrSerValIlyLeuCysThr 159

RESULT 5  
US-09-599-564A-2  
; Sequence 2, Application US/09599564A  
; Patent No. 6362318  
; GENERAL INFORMATION:  
; APPLICANT: KOMAN, Ahment  
; APPLICANT: CHASSIN, Dordine  
; APPLICANT: BELLET, Dominique  
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE  
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 017753-127  
; CURRENT APPLICATION NUMBER: US/09/599,564A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 09/174,465  
; PRIOR FILING DATE: 1998-10-19  
; PRIOR APPLICATION NUMBER: US 08/482,842  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early  
US-09-599-564A-2

Alignment Scores:  
Pred. No.: 2,02e-68 Length: 139  
Score: 613.00 Matches: 113  
Percent Similarity: 99.12% Conservative: 0  
Best Local Similarity: 99.12% Mismatches: 1  
Query Match: 96.54% Indels: 0  
Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-09-599-564A-2 (1-139)

QY 1 GCAGAGCTGAGGGGATGTGCTCCCGATTGGAAAAACACTTGTCTATATATGCCCCATG 60  
Db 26 AlaGluLeuArgIlyCysGlyProArgPheGlyLysHisLeuSerTyrCysPromet 45  
QY 61 CCTGAGAGACATTCACACACCCCGAGGGGCTGTGGAATCTGACGCTCCCAA 120  
Db 46 ProGluLysThrPheThrThrThrProGlyLyrProLeuLeuGluSerGIyArgProLys 65  
QY 121 GAATGTGTGTCACCTCCACAAACAAAGATGACAAACCTTAGGTACGACATTCGAATTC 180  
Db 66 GluMetValSerThrSerLysAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85  
QY 181 ATTCCTAATTTGTACACAGAGCTGAGAAACCACTGTCTGAGAGGCGACCATATTCAG 240  
Db 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerGIyGlnProSerLeuLys 105  
QY 241 AAAATATACTTCCCGAAAAAGAGAGAGTGCACAGATTGATTCATTCCTGTTGT 300  
Db 106 LysIleIleLeuSerArgIlyLysArgSerGIyArgHisArgPheAspProPheCysCys 125  
QY 301 GAAGTAATTGTGACGATGACGTAATTCATTAATATGATGACA 342  
Db 126 GluValIleIleCysAspAspGlyThrSerValIlyLeuCysThr 139

RESULT 6  
US-09-174-465D-2  
; Sequence 2, Application US/09174465D

Patent No. 6180364  
; GENERAL INFORMATION:  
; APPLICANT: KOMAN, Ahment  
; APPLICANT: CHASSIN, Dordine  
; APPLICANT: BELLET, Dominique  
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE  
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 017753-103  
; CURRENT APPLICATION NUMBER: US/09/174,465D  
; CURRENT FILING DATE: 1998-10-19  
; PRIOR APPLICATION NUMBER: US 08/482,842  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early  
US-09-174-465D-2

Alignment Scores:  
Pred. No.: 9.33e-68 Length: 174  
Score: 608.00 Matches: 113  
Percent Similarity: 99.12% Conservative: 0  
Best Local Similarity: 99.12% Mismatches: 1  
Query Match: 95.75% Indels: 0  
Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-09-174-465D-2 (1-174)

QY 1 GCAGAGCTGAGGGGATGTGCTCCCGATTGGAAAAACACTTGTCTATATATGCCCCATG 60  
Db 61 AlaGluLeuArgIlyCysGlyProArgPheGlyLysHisLeuSerTyrCysPromet 80  
QY 61 CCTGAGAGACATTCACACACCCCGAGGGGCTGTGGAATCTGACGCTCCCAA 120  
Db 81 ProGluLysThrPheThrThrThrProGlyLyrProLeuLeuGluSerGIyArgProLys 100  
QY 121 GAATGTGTGTCACCTCCACAAACAAAGATGACAAACCTTAGGTACGACATTCGAATTC 180  
Db 101 GluMetValSerThrSerLysAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 120  
QY 181 ATTCCTAATTTGTACACAGAGCTGAGAAACCACTGTCTGAGAGGCGACCATATTCAG 240  
Db 121 IleProAsnLeuSerProGluLeuLysLysProLeuSerGIyGlnProSerLeuLys 140  
QY 241 AAAATATACTTCCCGAAAAAGAGAGAGTGCACAGATTGATTCATTCCTGTTGT 300  
Db 141 LysIleIleLeuSerArgIlyLysArgSerGIyArgHisArgPheAspProPheCysCys 160  
QY 301 GAAGTAATTGTGACGATGACGTAATTCATTAATATGATGACA 342  
Db 161 GluValIleIleCysAspAspGlyThrSerValIlyLeuCysThr 174

RESULT 7  
US-09-174-465D-10  
; Sequence 10, Application US/09174465D  
; Patent No. 6180364  
; GENERAL INFORMATION:  
; APPLICANT: KOMAN, Ahment  
; APPLICANT: CHASSIN, Dordine  
; APPLICANT: BELLET, Dominique  
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE  
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID  
; FILE REFERENCE: 017753-103  
; CURRENT APPLICATION NUMBER: US/09/174,465D



```

; CURRENT FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like Peptide
US-09-174-465D-10

Alignment Scores:
Pred. No.: 8,36e-24 Length: 51
Score: 255.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.16% Indels: 0
DB: 4 Gaps: 0

US-09-518-842-1_COPY_76_417 (1-342) x US-09-174-465D-10 (1-51)
QY 100 CTGGAATCTGAGCTCCCAAGAAATGCTGTCACTCCACACACAAAGATGACAGCC 159
Db 1 LeuGlSerclYArProlyslumelValSerThSerAsnAsnlyAspGlylnAla 20
QY 160 TTAGGTACGACATCAGATTCATTCCTAATTGTGCACGAGGTGAAGAACCACTGTCT 219
Db 21 LeuGlYThrThrSerGluPheIleProAsnLeuSerProGluLeuYslsProLeuSer 40
QY 220 GAAGGCGACCATTCATTGAAGAAATTAATCTT 252
Db 41 GluGlyGlnProSerLeuYslsIleIleu 51

RESULT 8
US-09-599-564A-10
; Sequence 10, Application US/09599564A
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dordine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-127
; CURRENT APPLICATION NUMBER: US/09/599,564A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/174,465
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like Peptide
US-09-599-564A-10

Alignment Scores:
Pred. No.: 8,36e-24 Length: 51
Score: 255.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.16% Indels: 0
DB: 4 Gaps: 0
```

```

US-09-518-842-1_COPY_76_417 (1-342) x US-09-599-564A-10 (1-51)
QY 100 CTGGAATCTGAGCTCCCAAGAAATGCTGTCACTCCACACACAAAGATGACAGCC 159
Db 1 LeuGlSerclYArProlyslumelValSerThSerAsnAsnlyAspGlylnAla 20
QY 160 TTAGGTACGACATCAGATTCATTCCTAATTGTGCACGAGGTGAAGAACCACTGTCT 219
Db 21 LeuGlYThrThrSerGluPheIleProAsnLeuSerProGluLeuYslsProLeuSer 40
QY 220 GAAGGCGACCATTCATTGAAGAAATTAATCTT 252
Db 41 GluGlyGlnProSerLeuYslsIleIleu 51

RESULT 9
US-09-174-465D-8
; Sequence 8, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dordine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; CURRENT FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like Peptide
US-09-174-465D-8

Alignment Scores:
Pred. No.: 4,38e-16 Length: 41
Score: 193.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.39% Indels: 0
DB: 4 Gaps: 0

US-09-518-842-1_COPY_76_417 (1-342) x US-09-174-465D-8 (1-41)
QY 1 GCAAGCTGAGGAGATGTGCTCCCGATTTGGAACACATTCGTGTATATGCCCCATG 60
Db 9 AlaGluLeuArGlyCysGlyProArGpHeGlyshsIleuSerTyrcysPromet 28
QY 61 CCTGAGACATTCACACACACCCGAGGAGGGGTG 99
Db 29 ProGluYslsThrPheThrThrProGlyGlyTyrPleu 41

RESULT 10
US-09-599-564A-8
; Sequence 8, Application US/09599564A
; Patent No. 6362318
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dordine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
```

```
FILE REFERENCE: 017753-127
CURRENT APPLICATION NUMBER: US/09/599,564A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 09/174,465
PRIOR FILING DATE: 1998-10-19
PRIOR APPLICATION NUMBER: US 08/482,842
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 8
LENGTH: 41
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
US-09-599-564A-8

Alignment Scores:
Pred. No.: 4,38e-16 Length: 41
Score: 193.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.39% Indels: 0
DB: 4 Gaps: 0

US-09-518-842-1_COPY_76_417 (1-342) x US-09-599-564A-8 (1-41)

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTGGAACACTTGTCATATTCGCCCATG 60
Db 9 AlacIuLeuAtrgLyGcysGlyProArGpHeGlyHisLeuLeuSerTyrcysPromet 28
QY 61 CCTGAGAGACATTCACCCAGAGGGGTGCTG 99
Db 29 ProGluLysThrPheThrThrProGlyGlyTyrLeu 41

RESULT 11
US-09-201-227A-27
Sequence 27, Application US/09201227A
Patent No. 6468770
GENERAL INFORMATION:
APPLICANT: Keyes, Linda N.
APPLICANT: Doberstein, Stephen K.
APPLICANT: Buchman, Andrew R.
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF D. MELANOGASTER INSULIN-LIKE
TITLE OF INVENTION: GENES AND USES THEREOF
FILE REFERENCE: 7326-066
CURRENT APPLICATION NUMBER: US/09/201,227A
CURRENT FILING DATE: 1998-11-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patentln version 3.0
SEQ ID NO 27
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
US-09-201-227A-27

Alignment Scores:
Pred. No.: 4,38e-16 Length: 41
Score: 193.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.39% Indels: 0
DB: 4 Gaps: 0

US-09-518-842-1_COPY_76_417 (1-342) x US-09-201-227A-27 (1-41)

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTGGAACACTTGTCATATTCGCCCATG 60
Db 9 AlacIuLeuAtrgLyGcysGlyProArGpHeGlyHisLeuLeuSerTyrcysPromet 28
QY 61 CCTGAGAGACATTCACCCAGAGGGGTGCTG 99
Db 29 ProGluLysThrPheThrThrProGlyGlyTyrLeu 41
```

```
Db 29 ProGluLysThrPheThrThrProGlyGlyTyrLeu 41

RESULT 12
US-09-174-465D-12
Sequence 12, Application US/09174465D
Patent No. 6180364
GENERAL INFORMATION:
APPLICANT: KOMAN, Alment
APPLICANT: CHASSIN, Doriane
APPLICANT: BELLET, Dominique
TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 017753-103
CURRENT APPLICATION NUMBER: US/09/174,465D
CURRENT FILING DATE: 1998-10-19
PRIOR APPLICATION NUMBER: US 08/482,842
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12
LENGTH: 30
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
US-09-174-465D-12

Alignment Scores:
Pred. No.: 9,24e-13 Length: 30
Score: 166.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.14% Indels: 0
DB: 4 Gaps: 0

US-09-518-842-1_COPY_76_417 (1-342) x US-09-174-465D-12 (1-30)

QY 253 TCCCGGAAAGAGAGAGGAGCGTCAGATTCATTCCTGTTGAGTAATTTGT 312
Db 1 SerArgLysLysArgSerGlyArgHisArgPheAspProPheGcysGluValIleGys 20
QY 313 GACGATGAACTTCAGTTAAATTAATGTACA 342
Db 21 AspAspGlyTThrSerValLysLeuGlyThr 30

RESULT 13
US-09-599-564A-12
Sequence 12, Application US/09599564A
Patent No. 6362318
GENERAL INFORMATION:
APPLICANT: KOMAN, Alment
APPLICANT: CHASSIN, Doriane
APPLICANT: BELLET, Dominique
TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 017753-127
CURRENT APPLICATION NUMBER: US/09/599,564A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 09/174,465
PRIOR FILING DATE: 1998-10-19
PRIOR APPLICATION NUMBER: US 08/482,842
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12
LENGTH: 30
TYPE: PRT
ORGANISM: Unknown
```

```

FEATURE:
OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
OTHER INFORMATION: Placenta Insulin-Like Peptide
US-09-599-564A-12

```

```

Alignment Scores:
Pred. No.: 9.24e-13 Length: 30
Score: 166.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.14% Indels: 0
DB: 4 Gaps: 0

```

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-09-599-564A-12 (1-30)

```

QY 253 TCCCGCAAAAAGAGAGTGCACAGATTGATTCATTCGTGTGAGAAATTTGCT 312
Db 1 SerArgLysArgSerArgLysArgLysArgLysArgLysArgLysArgLysArgLys 20

```

```

QY 313 GACGATGCACTTCACTTAAATTATGTACA 342
Db 21 AspAspGlyThrSerValLysLeuGlyThr 30

```

RESULT 14

```

US-09-201-227A-28
Sequence 28, Application US/09201227A
Patent No. 6468770
GENERAL INFORMATION:
APPLICANT: Keyes, Linda N.
APPLICANT: Doberstein, Stephen K.
APPLICANT: Buchanan, Andrew R.
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF D. MELANOCASTER INSULIN-LIKE
FILE REFERENCE: 7326-066
CURRENT APPLICATION NUMBER: US/09/201,227A
CURRENT FILING DATE: 1998-11-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.0
SEQ ID NO 28
LENGTH: 25
TYPE: PRT
ORGANISM: Homo sapiens
US-09-201-227A-28

```

```

Alignment Scores:
Pred. No.: 2.05e-10 Length: 25
Score: 147.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.15% Indels: 0
DB: 4 Gaps: 0

```

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-09-201-227A-28 (1-25)

```

QY 268 AGTGAAGCTGCACAGATTGATTCATTCGTGTGAGAAATTTGAGCAAGTTCGA 327
Db 1 SerGlyArgHisArgPheAspProPheCysGluValIleCysAspAspGlyThrSer 20

```

```

QY 328 GTTAAATTATGTACA 342
Db 21 ValLysLeuGlyThr 25

```

RESULT 15

```

US-08-443-568B-12
Sequence 12, Application US/08443568B
GENERAL INFORMATION:
APPLICANT: Breese, Tim
APPLICANT: Hayenga, Kirk
APPLICANT: Rindersknecht, Ernst
APPLICANT: Vandien, Richard
APPLICANT: Daniel, Yansura
TITLE OF INVENTION: Process for Producing Relaxin

```

```

NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,568B
FILING DATE: 22-MAY-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/080,354
FILING DATE: 21-JUNE-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B.
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 7842-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE

```

```

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-443-568B-12

```

```

Alignment Scores:
Pred. No.: 6.49e-05 Length: 150
Score: 105.50 Matches: 43
Percent Similarity: 35.67% Conservative: 13
Best Local Similarity: 27.39% Mismatches: 44
Query Match: 16.61% Indels: 57
DB: 1 Gaps: 6

```

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-08-443-568B-12 (1-150)

```

QY 1 GCAGAGCTGAGGGGATGTGT---CCCGATTGGAAAACACTTGTGTCATATGCCCC 57
Db 7 AlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGluAsp 26

```

```

QY 58 ATGCGTGAAGAAACATTCACACACACCCACAGAGGGGCTGCTGGAATCTGACGTCCC 117
Db 27 AlaProGln-----ThrPro-----ArgPro 33

```

```

QY 118 ---AAGAAATGTGTGTCACCTCCACACAAAGATGAGCAAGCTTAGTACAGATCA 174
Db 34 ValAlaGlnIleValAlaProSerPheIleAsnLysAspThrGlnThrIleAsnMetSer 53

```

```

QY 175 GAATTCATTCTCTTAATTGTGCACAGAGCTGAGAAACCACTGTCTGAAGGCGCCATCA 234
Db 54 GluPheValAlaAsnLeuProGlnGlnLeuLysLeuThrLeuSerGlnMetGlnProAla 73

```

```

QY 235 TTG----- 237
Db 74 LeuProGlnLeuGlnGlnIleHisValProValLeuLysAspSerSerLeuLeuPheGlnGlu 93

```

```

QY 238 ---AAGAAATTAATCTT----- 252
Db 94 PheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerSerProSerGlnLeu 113

```

```

QY 253 -----TCCCGCAAAAAGAGAGTGCACAGATTGAT 288

```

Db 114 LysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaLeuAla 133  
 QY 289 CCATCTGTTGAGAGTAATTGTGACGATGGAACCTTCAGTTAAATTATGCT 339  
 ||||| ||| |||  
 Db 134 AsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 150  
 :::  
 :::  
 |||

Search completed: June 14, 2003, 19:47:19  
 Job time : 15.5 secs

US-09-518-842-1\_COPY\_76\_417 (1-342) X US-10-007-693-98 (1-1531)

```

QY      4  GACCTGAGGGAGTGGCTCCCGCATTTGGAAACACTGG-----42
Db      472  GltuYrGlnIleYelYglYlAlaLeuPheGclYlUsAnIleSerLeuSerGluAnSlaGly 491
QY      43  --CTGTCAATTGCCCCATGCGCTGAGAAGACATTCCACCACC-----CCA 87
Db      492  ValIeuThrPheYlSAspAsnIleValIylStrPheAlaSerAsnGlyLysIleGluGly 511
QY      88  GAGGGGTGGTGGTGAATCTGACGCTCCAAAGAAGATGTGTGAACCTCCACACAA 147
Db      512  GlYclYAlaIleLeuAlaIhrGlyLys-----ValGluIleThrAsnAsnSer 527
QY      148  GATGGA-----CAAGCTTAGGTAGCATTCAGAA 177
Db      528  GluGlYlYleSerPhehIrglYsnaIaargAlaProGlnAlaIeProThGlnGluGlu 547
QY      178  TTCATTCCATTATTTGCACACAGCTGAGGAACCACTGTCGAAGGG 225
Db      548  Phe---ProlLeuPheSerLysLysGluGlyAlaGProLeuSerSerGly 562

```

[illegible]

Alignment Scores:		
Pred. No.:	4.46	Length: 269
Score:	73.50	Matches: 30
Percent Similarity:	44.23%	Conservative: 16
Best Local Similarity:	28.85%	Mismatches: 30
Query Match:	11.57%	Indels: 28
DB:	10	Gaps: 5

US-09-518-842-1\_COPY\_76\_417 (1-342) X US-09-815-242-5429 (1-269)

1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTGGAAAAACACTGCTGTCATATTGGCCCATG 60  
 ||| ::| ||| |||::| |||||::| |||

[illegible]

```

RESULT 3
US-09-815-242-12487
; Sequence 12487, Application US/09815242
; Patent No. US20020061559A1
;
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trivick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA-011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 12487
; LENGTH: 273
; TYPE: PRT
;
ORGANISM: Staphylococcus aureus
US-09-815-242-12487

```

Alignment Scores:	
Pred. No.:	4.48
Score:	73.50
Percent Similarity:	44.23%
Best Local Similarity:	28.85%
Query Match:	11.57%
DB:	10
Gaps:	5
Length:	273
Matches:	30
Conservative:	16
Mismatches:	38
Indels:	28

US-09-518-842-1\_COPY\_76\_417 (1-342) X US-09-815-242-12487 (1-273)

1 GCAGAGCTGAGGGGATGTGGTCCCGATTTCGAAACACTTGTCTTCATATTGCCCATG 60



```

Db      253 nh1stYrllleargAsnAlaHis-----LysArgGluLeu 264
QY      271 GGAGCTCACA 280
Db      264 uAlaTyrThr 267

RESULT 6
US-09-925-637-12
; Sequence 12, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Chai
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-637-12

Alignment Scores:
Pred. No.: 4.48 Length: 274
Score: 73.50 Matches: 30
Percent Similarity: 44.23% Conservative: 16
Best Local Similarity: 28.85% Mismatches: 30
Query Match: 11.57% Indels: 28
DB: Gaps: 5

US-09-518-842-1_COPY_76_417 (1-342) x US-09-925-637-12 (1-274)
QY      1 GCAGAGCTGAGGGAGTGGTCCCGATTGGAAACACTTGCCTGATATTGGCCCATG 60
Db      182 AlaprollephneglYLeuAsnLysArgGlnGlnLeuAlaTyrLeuGlyAla 201
QY      61 CCTGAGAGACATTCACACCCACCACCCAGAGGTGGCTGGAATCTGGACGTCCCAA 120
Db      202 ProlysgluLeuYrGlnLysThrProThrAla---AspleuGluAspLysProGln 220
QY      121 GAATGTGTCAACCTCCAAACAACAAGATGACACAGCTTAGTACGACATCGAATTC 180
Db      221 -----LeuProAspLysAlaLeuGlyValThrTyrGluAla 233
QY      181 ATTCTCAT-----TTGTCCACCAGAG-----CTGAGAAA 210
Db      234 IleAspAsnTyrLeuGlnGlyLysProValThrProGlnGlnGlnValIleGlu-As 253
QY      211 CCACGTGTGAAGGGCAGCCATCATTTGAAGAAATAATACTTCCCGCAAAAGAGAGT 270
Db      253 nh1stYrllleargAsnAlaHis-----LysArgGluLeu 264
QY      271 GGAGCTCACA 280
Db      264 uAlaTyrThr 267

RESULT 7
US-10-028-056-3
; Sequence 3, Application US/10028056
; Patent No. US20020152483A1

```

```

; GENERAL INFORMATION:
; APPLICANT: REUE, KAREN
; APPLICANT: PTEREV, MIKLOS
; TITLE OF INVENTION: A NOVEL GENE ASSOCIATED WITH REGULATION OF ADIPOSITV AND INSUL
; FILE REFERENCE: 4077-898010US
; CURRENT APPLICATION NUMBER: US/10/028,056
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,772
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-056-3

Alignment Scores:
Pred. No.: 6.56 Length: 890
Score: 73.50 Matches: 20
Percent Similarity: 47.06% Conservative: 12
Best Local Similarity: 29.41% Mismatches: 35
Query Match: 11.57% Indels: 1
DB: Gaps: 1

US-09-518-842-1_COPY_76_417 (1-342) x US-10-028-056-3 (1-890)
QY      79 ACCACCCAGAGGGGTGCTGGAATCTGGACGTCCCAAGAATGTGTCAACCTCC 138
Db      324 ThrLeuValGlyGlyAlaLeuLeuAspGlnAsnLysProGlnThrGluMetGlnPheVal 343
QY      139 AACACAAAGATGACCAACCCCTTAGGTACGACATCGAATTCATCTTAATTTGCACA 198
Db      344 AsnGlnGluAspLeuGlnThrLeuGlyAlaAlaIleProLeuLeuPrometIleGlu--- 362
QY      199 GAGCTGAGAAACCACTGTCTGAGAGGACGCCATCATTTGAACAATAATATCTTCCGC 258
Db      363 GluLeuLysProProSerAlaSerValAlaGlnThrAlaAsnLysThrAspSerProSer 382
QY      259 AAAAGAGAGAGTGGACGTGCACAGA 282
Db      383 ArgLysArgAspLysArgSerArg 390

RESULT 8
US-09-841-132-193
; Sequence 193, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhalla, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 193
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-193

Alignment Scores:
Pred. No.: 11 Length: 778
Score: 71.50 Matches: 29
Percent Similarity: 43.75% Conservative: 13
Best Local Similarity: 30.21% Mismatches: 27
Query Match: 11.26% Indels: 27
DB: Gaps: 5

US-09-518-842-1_COPY_76_417 (1-342) x US-09-841-132-193 (1-778)

```







SOFTWARE: Patentin version 3.0  
SEQ ID NO 2  
LENGTH: 952  
TYPE: PRT  
ORGANISM: BACILLUS CALDOTEANX  
US-10-216-556-2

## Alignment Scores:

Pred. No.:	31	Length:	952
Score:	68.00	Matches:	20
Percent Similarity:	46.15%	Conservative:	10
Best Local Similarity:	30.77%	Mismatches:	17
Query Match:	10.71%	Indels:	18
DB:	9	Gaps:	3

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-10-216-556-2 (1-952)

```
QY 19 GGTCCCGCATTTGGAAACACTTGTGTCATATTGCCCATGCCCTGAGAGACATTCACCC 78
    |||||
DB 559 GYProGlyAlaGlyIleHis----- 565
QY 79 ACCACCCGAGGAGGTGCTGGAATCTGACGCTCCCAAGAAATGGTCAACCTCC 138
    |||||
DB 566 -----GlyGlyGluValIleSerAlaGlyThrProGluGluValMetGluAspPro 582
QY 139 AACACAAAGATGACACAGCCTTAGTACGACATCAGAAATTCATTCCTAAATTGTGCACCA 198
    |||||
DB 583 AsnSerLeuThrGlySerTyLeu---SerGlyLysLysPheIlePro---LeuProPro 600
QY 199 GAGCTGAAGAACA 213
    |||
DB 601 GluArgAlaGlyLysPro 605
```

## RESULT 15

US-09-738-626-5066  
Sequence 5066, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: Patentin ver. 3.0  
SEQ ID NO 5066  
LENGTH: 554  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5066

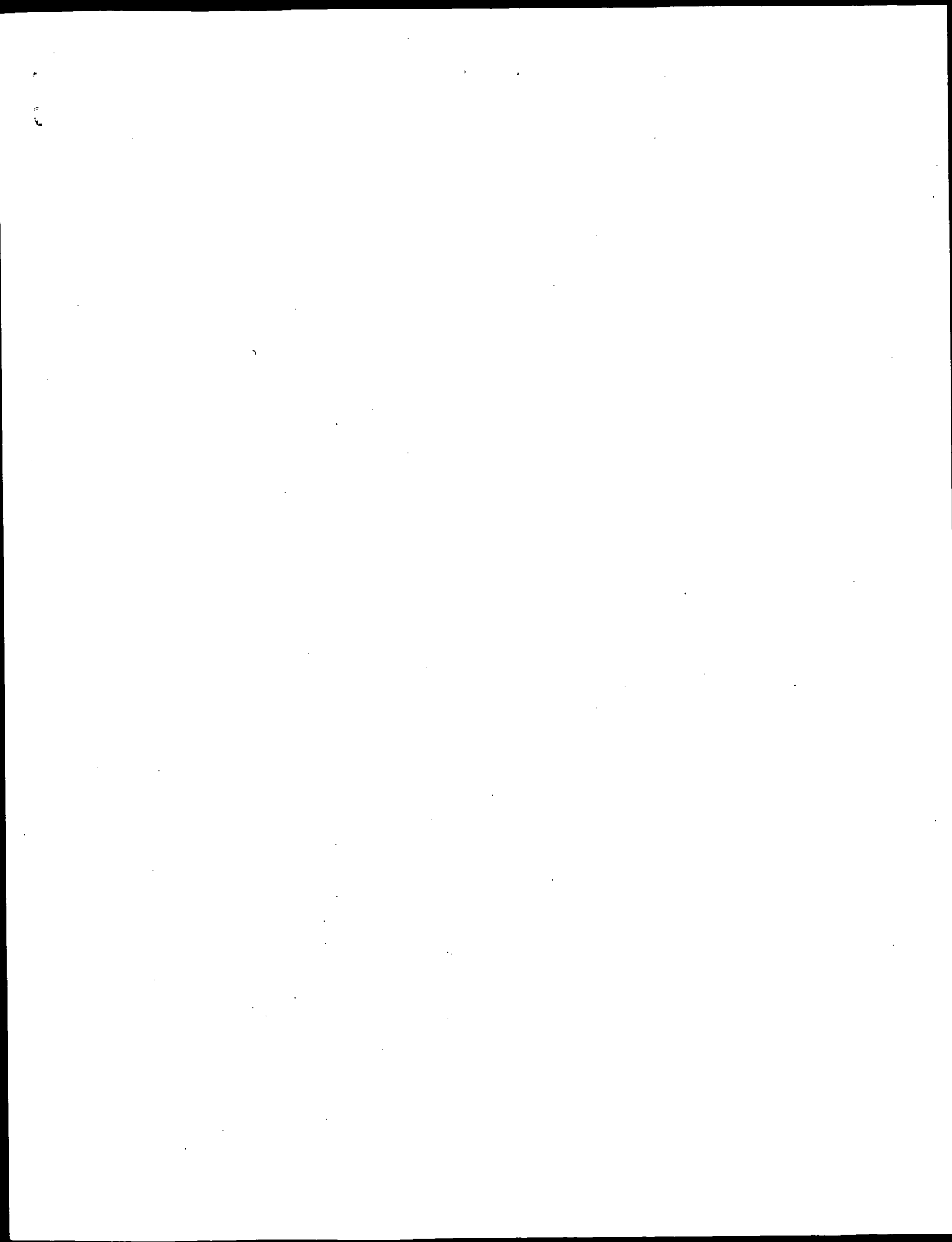
## Alignment Scores:

Pred. No.:	30	Length:	554
Score:	67.50	Matches:	16
Percent Similarity:	51.79%	Conservative:	13
Best Local Similarity:	28.57%	Mismatches:	24
Query Match:	11.25%	Indels:	3
DB:	9	Gaps:	2

US-09-518-842-1\_COPY\_76\_417 (1-342) x US-09-738-626-5066 (1-554)

```
QY 197 GGTGACAAATTAGGAATGAATTCGATGCTGACTAAGGCTGTCATCTTGTGTG 138
    |||||
DB 113 GlyIuTyIleuGlyLysThrValGluValIleProHisIleThrAspGluIleLysAla 132
QY 137 GAGTTGACACCACTTTCTTTGGAGCTCCAGATTCACGACGCCCTCGGGTGTG 78
    |||||
DB 133 ArgIle-----LeuSerMetGlyGluProAspAlaHisGlyAsnAlaProAspValVal 150
QY 77 GTGAATGCTTTCACAGC---ATGGGCAATATGACAGCAAGCTTT 33
    |||||
DB 151 ILeSerGluValGlyGlyThrValGlyAspIleGluSerGluProPhe 166
```

Search completed: June 14, 2003, 19:58:26  
Job time: 39 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK, nucleic - protein search, using frame\_plus\_n2p model

Run on: June 14, 2003, 19:30:20 ; Search time 34 seconds

(without alignments)  
2680.689 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 635

Sequence: 1 GCAGAGCTGAGGGAGGTGG.....CTTCAGTTAATATGATACA 342

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_n2p\_model -DEV=xlp  
-O=/cgn2\_1/USPFO.spool/US09518842/rnat\_14062003\_175810\_12235/app.query.fasta\_1\_519  
-DB=A\_Geneseq\_101002 -OPMT=fastan -SUFFIX=n2p\_rag -NIMATCH=0.1 -LOOPT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human+0.csi  
-LIST=45 -DOCALLIG=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MOTF=LOCAL -OUTPMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09518842\_8CGN\_1\_114\_8rnat\_14062003\_175810\_12235 -NCPU=6 -ICPU=3  
-NO\_MMAP -JARQUEURY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
A\_Geneseq\_101002 :  
1: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1980.DAT : \*  
2: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1981.DAT : \*  
3: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1982.DAT : \*  
4: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1983.DAT : \*  
5: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1984.DAT : \*  
6: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1985.DAT : \*  
7: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1986.DAT : \*  
8: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1987.DAT : \*  
9: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1988.DAT : \*  
10: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1989.DAT : \*  
11: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1990.DAT : \*  
12: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1991.DAT : \*  
13: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1992.DAT : \*  
14: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1993.DAT : \*  
15: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1994.DAT : \*  
16: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1995.DAT : \*  
17: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1996.DAT : \*  
18: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1997.DAT : \*  
19: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1998.DAT : \*  
20: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA1999.DAT : \*  
21: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA2000.DAT : \*  
22: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA2001.DAT : \*  
23: /SIDS2/gcgdata/geneseq/geneseq\_emb1/AA2002.DAT : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	619	97.5	124	19	AAW69169
2	619	97.5	139	17	AAW69134
3	619	97.5	139	18	AAW17676
4	619	97.5	139	19	AAW69168
5	619	97.5	139	20	AAW26926
6	619	97.5	139	20	AAW95574
7	619	97.5	159	19	AAW69170
8	142	22.4	182	5	AAW40156
9	141	22.2	182	4	AAW30392
10	111	17.5	178	14	AAW31958
11	105.5	16.6	150	16	AAW64904
12	105.5	16.6	162	11	AAW07987
13	105.5	16.6	185	10	AAW94621
14	103.5	16.3	185	5	AAW40154
15	103.5	16.3	185	5	AAW40155
16	103.5	16.3	185	10	AAW94622
17	103.5	16.3	185	23	AAW61819
18	103	16.2	185	5	AAW40108
19	103	17.2	787	22	ABG08971
20	103	17.2	787	22	ABG08971
21	99.5	15.7	164	11	AAW07988
22	97	15.3	220	22	AAW02911
23	78	12.3	927	22	ABG25847
24	75.5	11.9	612	22	AAW61791
25	74.5	11.7	1531	22	AAW38904
26	73.5	11.6	269	22	AAU33933
27	73.5	11.6	273	22	AAU36894
28	73.5	11.6	273	22	AAU37154
29	73.5	11.6	274	22	AAU00833
30	73.5	11.6	282	22	AAW67271
31	73	12.2	402	21	AAW42944
32	73	12.2	210	22	AAW49608
33	73	11.5	545	23	ABW89399
34	73	11.5	636	22	AAW94352
35	73	11.5	747	22	AAW40070
36	73	11.5	747	22	AAW93124
37	73	12.2	994	18	AAW27147
38	72.5	12.1	1754	22	ABG18818
39	72.5	12.1	1759	22	AAW09273
40	72	11.3	732	20	AAW42697
41	72	11.3	806	22	AAW65622
42	71.5	11.3	647	23	ABP26017
43	71.5	11.3	778	21	AAW13642
44	71.5	11.3	778	22	AAW83210
45	71.5	11.3	778	23	ABW94181

## ALIGNMENTS

RESULT 1  
ID AAW69169 standard; Protein: 124 AA.

AAW69169; 07-OCT-1998 (first entry)

DE Zins1 NF protein.

KW Zins1; human; placenta; pancreatic islet cell proliferation;

KW insulin secretion; diabetes; therapy; Zins1 NF.

OS Homo sapiens.

PN W09827210-A1.

PD 25-JUN-1998.

PF 16-DEC-1997; 97WO-US23326.  
XX  
XX 16-DEC-1996; 96US-0033003.  
XX  
XX (ZYMO) ZYMOGENETICS INC.  
XX  
XX Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;  
PI Sprugel KH;  
XX  
XX WPI; 1998-362779/31.  
XX  
XX New isolated protein, Zinsl - obtained from human placenta, which  
PT increases the proliferation of pancreatic islet cells, used for  
PT treating diabetes  
XX  
XX Example 3; Page 64; 77pp; English.  
XX  
XX This sequence is the human Zinsl NF protein of the invention. The Zinsl  
CC protein was isolated from human placenta, and is believed to be a new  
CC version of the mature protein of placenta, having disulphide bonded A  
CC and B chains. The protein can be used for stimulating the proliferation  
CC of pancreatic islets to increase insulin secretory capacity of mammals.  
CC In particular it can be used for the treatment of diabetes. It can also  
CC be used for stimulating in vitro proliferation of pancreatic islet cells.  
CC It can also be used for production of antibodies and in detection and  
CC diagnosis.  
XX  
XX Sequence 124 AA:  
SQ  
Alignment Scores:  
Pred. No.: 3.7e-68 Length: 124  
Score: 619.00 Matches: 114  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.48% Indels: 0  
DB: 19 Gaps: 0  
US-09-518-842-1\_COPY\_76\_417 (1-342) x AAW69169 (1-124)  
QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTTGAAAAACCTTGCTCATATTGCCCATG 60  
DB 11 AlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuSerTyrCysPromet 30  
QY 61 CCTGAGAGACATTCACACACACCCAGAGGGGTGCTGGAATCTGGACGTCCCAA 120  
DB 31 ProGluLysThrPheThrThrThrProGlyLysPheLeuGluSerGlyArgProlys 50  
QY 121 GAATGGTGTCAACCTCCCAACAAGAAGTGGACAGCCTTGTGATCAGATCAGAAATTC 180  
DB 51 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 70  
QY 181 ATTCCTAAATTTGACACAGAGCTGAAGAACAACACTGTGAGAGGACAGCCATCATTTGAG 240  
DB 71 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlnProSerLeuLys 90  
QY 241 AAAATAACTCTTCCCGCAAAAAGAGAGAGTGCAGCTGACAGATTTGATTCATTCTGTGT 300  
DB 91 LysIleIleLeuSerArgLysLysArgSerGlyArgHisArgPheAspProPheCysCys 110  
QY 301 GAAGTAATTTGACAGCTGAGATGAGTAAATTAATGATACA 342  
DB 111 GluValIleCysAspAspGlyThrSerValLysLeuCysThr 124  
RESULT 2  
AAR89134  
ID AAR89134 standard; protein; 139 AA.  
AC AAR89134;  
XX  
XX 31-JUL-1996 (first entry)  
XX  
XX Human early placental insulin-like protein.  
XX

KW Insulin growth hormone family; early placental insulin like protein;  
KW placental; placental tissue; cytotrophoblast; trimester; probe; primer;  
KW amplification; polymerase chain reaction; tyrosine phosphorylation;  
KW cellular protein; growth factor; human; lactation; promoter; PCR;  
KW regeneration; nerve; muscle; skin; bone tissue.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..17  
FT Peptide /note="signal peptide"  
FT Peptide 18..58  
FT Peptide /note="B-chain peptide"  
FT Peptide 59..109  
FT Peptide /note="C-peptide, links B and A chains in pre-E"  
FT Peptide 110..139  
FT Peptide /note="A-chain peptide"  
XX  
XX W09534653-A1.  
XX  
XX 21-DEC-1995.  
XX  
XX 12-JUN-1995; 95WO-FR00766.  
XX  
XX 13-JUN-1994; 94FR-0007191.  
XX  
XX (INSR) INST ROUSSY GUSTAVE.  
XX  
XX Bellet D, Chassin D, Koman A;  
XX  
XX WPI; 1996-049682/05.  
XX  
XX N-PSDB; AAT10275.  
XX  
XX Early placental insulin-like protein, EPIL/placental - contains  
PT growth factor-like activity useful for e.g. promoting lactation or  
PT for regeneration of nerve, muscle, skin or bone tissue  
XX  
XX  
XX Claim 1; Page 14-15; 25pp; French.  
XX  
XX  
XX This is the amino acid sequence of a novel member of the insulin growth  
CC hormone family designated early placental insulin like (EPIL) protein or  
CC placental. The encoding gene has been found to be expressed exclusively  
CC in the early placental tissue with a small amount in full term placental  
CC tissue. No other tissue expresses this gene. The gene was obtained from  
CC a cDNA library derived from mRNA isolated from cytotrophoblasts derived  
CC from first trimester placental tissue using, as a probe, a fragment of  
CC the gene amplified by primers AAT10276-7. Although the specific  
CC activity of the protein remains to be elucidated, it is thought that the  
CC protein will induce tyrosine phosphorylation of cellular proteins and may  
CC have growth factor-like activities e.g. human growth factor type 1 or  
CC lactation promoter activities. It may also be used to regenerate e.g.  
CC nerve, muscle, skin or bone tissue.  
XX  
XX Sequence 139 AA:  
SQ  
Alignment Scores:  
Pred. No.: 3.85e-68 Length: 139  
Score: 619.00 Matches: 114  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.48% Indels: 0  
DB: 17 Gaps: 0  
US-09-518-842-1\_COPY\_76\_417 (1-342) x AAR89134 (1-139)  
QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTTGAAAAACCTTGCTCATATTGCCCATG 60  
DB 26 AlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuSerTyrCysPromet 45  
QY 61 CCTGAGAGACATTCACACACACCCAGAGGGGTGCTGGAATCTGGACGTCCCAA 120  
DB 46 ProGluLysThrPheThrThrThrProGlyLysPheLeuGluSerGlyArgProlys 65

```

QY 121 GAAATGCTGCACCTCCACCAACAAAGATGACAGACCTTAGTACGACATCAGAAATTC 180
    |||||||
Db 66 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85
    |||||||
QY 181 ATTCCCTAATTGTCACAGAGCGTGAAGAAACCACTGCTCTGAAGGACCCATCATTTGAG 240
    |||||||
Db 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerIleuGlyInProSerLeuLys 105
    |||||||
QY 241 AAAATAATCTCTCCGCCAAAGAGAGAGTGCAGTGCACAGATTGATTCATCTGTTGT 300
    |||||||
Db 106 LysIleIleLeuSerArgLysLysArgSerGlyArgHisArgPheAspProPheCys 125
    |||||||
QY 301 GAAGTAATTTGTGACAGATGAGACTGATTAATTAATGTACA 342
    |||||||
Db 126 GluValIleCysAspAspGlyThrSerValLysLeuCysThr 139
    |||||||
RESULT 3
AAW17676
ID AAW17676 standard; Protein; 139 AA.
XX
AC AAW17676;
XX
DT 24-JUL-1997 (first entry)
XX
DE Human relaxin-related factor-2 (RRF-2).
XX
KW Relaxin-related factor-2; RRF-2; testis; sperm; growth factor;
    infertility.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..26
FT Misc-difference 124 /note= "conserved Cys residue indicative of
    insulin family member"
FT Misc-difference 125 /note= "conserved Cys residue indicative of
    insulin family member"
FT Misc-difference 129 /note= "conserved Cys residue indicative of
    insulin family member"
FT Misc-difference 138 /note= "conserved Cys residue indicative of
    insulin family member"
FT Misc-difference 138 /note= "conserved Cys residue indicative of
    insulin family member"
XX
PN W09716549-A2.
XX
PD 09-MAY-1997.
XX
PF 01-NOV-1996; 96WO-US17342.
XX
PR 21-FEB-1996; 96US-0012016.
    03-NOV-1995; 96US-0006221.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Davis S.
XX
DR MPI; 1997-272118/24.
    N-PSDB; AAT68419.
XX
PT New isolated relaxin-related factor genes - used to develop products
    which can be used in diagnosis and therapy, e.g. in fertility and
    pregnancy applications
XX
PS Example 3; Fig 4; 34pp; English.
XX
CC Human relaxin-related factor-2 (RRF-2) (AAW17676) is a placenta-
    specific growth factor related to relaxin and to the insulin family
    of ligands. RRF-2 cDNA (AAT68419) was isolated in a search of
    expressed sequence tags for sequences related to relaxin. RRF-2

```

```

CC displays all the expected features of a new insulin family member,
CC partic. with regard to a cluster of four cysteine residues at the
CC C-terminus of the molecule. A related testis-specific factor,
CC RRF-1 (AAW17675), has also been identified. RRF-2 can be produced
CC in transformed host cells for use in the prepn. of antibodies and
CC therapeutic compns., or as a growth factor for maintaining cells
CC in culture. RRF-2 may be useful for modulating the reproductive
    physiology of mammals during pregnancy and parturition.
XX
SQ Sequence 139 AA:
Alignment Scores:
Pred. No.: 3 85e-68 Length: 139
Score: 619.00 Matches: 114
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.48% Indels: 0
DB: 18 Gaps: 0
US-09-518-842-1_COPY_76_417 (1-342) x AAW17676 (1-139)
QY 1 GCAGAGCTGAGGGGAGATGTGTCGCCGATTTGAAACACTTGCTCATATTGCCCATG 60
    |||||||
Db 26 AlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuSerTyrCysPromet 45
    |||||||
QY 61 CCTGAGAGACATTCACACACCCAGAGAGGTCGCTGGAATCTGGACGTCCCAA 120
    |||||||
Db 46 ProGluLysThrPheThrThrThrProGlyLysPheLeuGluGlnSerGlyArgProLys 65
    |||||||
QY 121 GAAATGCTGCACCTCCACCAACAAAGATGACAGACCTTAGTACGACATCAGAAATTC 180
    |||||||
Db 66 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85
    |||||||
QY 181 ATTCCCTAATTGTCACAGAGCTGAAGAAACCACTGCTCTGAAGGACCCATCATTTGAG 240
    |||||||
Db 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerIleuGlyInProSerLeuLys 105
    |||||||
QY 241 AAAATAATCTCTCCGCCAAAGAGAGAGTGCAGTGCACAGATTGATTCATCTGTTGT 300
    |||||||
Db 106 LysIleIleLeuSerArgLysLysArgSerGlyArgHisArgPheAspProPheCys 125
    |||||||
QY 301 GAAGTAATTTGTGACAGATGAGACTGATTAATTAATGTACA 342
    |||||||
Db 126 GluValIleCysAspAspGlyThrSerValLysLeuCysThr 139
    |||||||
RESULT 4
AAW69168
ID AAW69168 standard; Protein; 139 AA.
XX
AC AAW69168;
XX
DT 07-OCT-1998 (first entry)
XX
DE Zinsl protein.
XX
KW Zinsl; human; placenta; placental; pancreatic islet cell proliferation;
    insulin secretion; diabetes; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 26..43
FT Region /note= "B chain of Zinsl"
FT Region 115..0
FT Region /note= "A chain of Zinsl"
XX
PN W09827210-A1.
XX
PD 25-JUN-1998.
XX
PF 16-DEC-1997; 97WO-US23326.
    16-DEC-1996; 96US-0033003.
PR

```

XX (ZYMO ) ZYMOGENETICS INC.  
 XX Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP:  
 PI Sprugel KH:  
 XX WPI: 1998-362779/31.  
 DR N-PSDB: AAV44663.  
 XX  
 PT New isolated protein, Zinsl - obtained from human placenta, which  
 PT increases the proliferation of pancreatic islet cells, used for  
 PT treating diabetes  
 XX  
 PS Claim 4; Page 60-61; 77pp; English.  
 XX  
 CC This sequence is the human zinsl protein of the invention. The zinsl  
 CC protein was isolated from human placenta, and is believed to be a new  
 CC version of the mature protein of placentalin, having disulphide bonded A  
 CC and B chains. The protein can be used for stimulating the proliferation  
 CC of pancreatic islets to increase insulin secretory capacity of mammals.  
 CC In particular it can be used for the treatment of diabetes. It can also  
 CC be used for stimulating in vitro proliferation of pancreatic islet cells.  
 CC It can also be used for production of antibodies and in detection and  
 CC diagnosis.  
 XX  
 SO Sequence 139 AA:  
 Alignment Scores:  
 Pred. No.: 3,85e-68 Length: 139  
 Score: 619.00 Matches: 114  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.48% Indels: 0  
 DB: Gaps: 0  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x AAM69168 (1-139)  
 QY 1 GCAGAGCTGAGGGATGTCGCCGATTTGGAACACTTGTCTCATATTGCCCCATG 60  
 DB 26 AlaGluLeuArgGlyGlySerGlyProArgPheGlyLysHisLeuSerTyrCysPromet 45  
 QY 61 CCTGAGAGACATTCACCCACCCAGAGAGGCTGCTGCAATCTGCACGTCCTCCAA 120  
 DB 46 ProGluLysThrPheThrThrThrProGlyLysTyrLeuGluSerGlyArgProLys 65  
 QY 121 GAATGCTGTCACCTCCACCAACAAGATGAGCAAGCTTACGATCAGATTCGAAATTC 180  
 DB 66 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85  
 QY 181 ATTCTAATTTGTACACAGAGCTGAGAAACCACTGCTCGAAGGCGACCATTCATTGAAG 240  
 DB 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlnProSerLeuLys 105  
 QY 241 AAAATAATACTTTCCCGCAAAAAGAGAGTGCAGATTCATTCCTGCTGCT 300  
 DB 106 LysIleIleLeuSerArgLysLysArgSerIleArgHisArgPheAspProPheCysCys 125  
 QY 301 GAAGTAATTTGTGACGATGGAAGCTCACTTAATTAATGATGATCA 342  
 DB 126 GluValIleCysAspAspGlyThrSerValLysLeuGlySerThr 139  
 RESULT 5  
 AAY26926  
 ID AAY26926 standard; Protein; 139 AA.  
 XX  
 AC AAY26926;  
 DT 21-DEC-1999 (first entry)  
 XX  
 DE Human insulin-like 4 protein.  
 XX  
 KW Human; insulin-like 4; INS14; embryonic; c bone tissue; ligament;  
 KW early-placental insulin-like protein; EPII; antibody; probe; primer;

KW diagnosis; pathology; differentiation; proliferation; cartilage;  
 KW ossification; osteoporosis; dysplasia.  
 XX Homo sapiens.  
 XX MO9937780-A1.  
 XX  
 XX 29-JUL-1999.  
 XX  
 XX 22-JAN-1999; 99WO-FR00137.  
 XX  
 XX 23-JAN-1998; 98FR-0000715.  
 XX  
 XX (INSR ) INST ROUSSY GUSTAVE.  
 XX  
 XX Laurent A, Bellet D:  
 XX WPI: 1999-469135/39.  
 XX DR N-PSDB: AAX83561.  
 XX  
 PT Expression of the INS14 gene in human embryonic bone tissue and  
 PT ligaments,  
 XX  
 PS Disclosure: Fig 1; 60pp; French.  
 XX  
 CC This sequence represents the human protein encoded by the insulin-like 4  
 CC (INS14) gene which is designated early-placental insulin-like (EPII)  
 CC protein. The INS14 gene is expressed in human embryonic bone tissue and  
 CC ligaments and encodes 3 different EPII proteins designated EPII 1, 2 or  
 CC 3. EPII 1 is a single chain comprising amino acids 115-139, EPII 2 is a  
 CC 2 chain protein with chain A comprising amino acids 18-139 and chain B  
 CC comprising amino acids 18-58 and EPII 3 is a single chain comprising  
 CC amino acids 59-114. The nucleic acids and protein or antibodies against  
 CC EPII 1, 2 or 3, and probes or primers for INS14 are useful for the  
 CC diagnosis of pathology associated with abnormal differentiation and/or  
 CC proliferation of bone tissue or ligaments or development of abnormal  
 CC cartilage and/or abnormal ossification of forming bones, e.g.  
 CC osteoporosis or dysplasia. Compositions capable of modulation  
 CC differentiation, regeneration and/or proliferation of bone tissue  
 CC and/or ligament cells are also useful in treating bone disease.  
 XX  
 SO Sequence 139 AA:  
 Alignment Scores:  
 Pred. No.: 3,85e-68 Length: 139  
 Score: 619.00 Matches: 114  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.48% Indels: 0  
 DB: Gaps: 0  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x AAY26926 (1-139)  
 QY 1 GCAGAGCTGAGGGATGTCGCCGATTTGGAACACTTGTCTCATATTGCCCCATG 60  
 DB 26 AlaGluLeuArgGlyGlySerGlyProArgPheGlyLysHisLeuSerTyrCysPromet 45  
 QY 61 CCTGAGAGACATTCACCCACCCAGAGAGGCTGCTGCAATCTGCAGTCCCAAA 120  
 DB 46 ProGluLysThrPheThrThrThrProGlyLysTyrLeuGluSerGlyArgProLys 65  
 QY 121 GAATGCTGTCACCTCCACCAACAAGATGAGCAAGCTTACGATCAGATTCGAAATTC 180  
 DB 66 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 85  
 QY 181 ATTCTAATTTGTACACAGAGCTGAGAAACCACTGCTCGAAGGCGACCATTCATTGAAG 240  
 DB 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlnProSerLeuLys 105  
 QY 241 AAAATAATACTTTCCCGCAAAAAGAGAGTGCAGATTCATTCCTGCTGCT 300  
 DB 106 LysIleIleLeuSerArgLysLysArgSerIleArgHisArgPheAspProPheCysCys 125



QY 301 GAAGTAATTTGTGACGATGGAAGTCTAGTTAATTATGTACA 342  
 |||||||  
 Db 126 GltuValIleCysaspaspGlyThrSerValIysLeucCysThr 139

RESULT 6  
 AAW99574  
 ID AAW99574 standard; Protein; 139 AA.  
 XX  
 AC AAW99574;  
 XX

DT 22-JUN-1999 (first entry)  
 XX

DE Human early placental insulin-like (EPIL) polypeptide.

KW INSU-4; insulin-like gene; EPIL; early placental insulin-like; antibody;  
 KM vector; probe; hybridisation; tumour; hypoglycaemia; hyperglycaemia;  
 XX diabetes; cardiovascular.

OS Homo sapiens.

PN WO9909172-A1.

PD 25-FEB-1999.

PF 12-AUG-1998; 98WO-FR01799.

PR 03-NOV-1997; 97FR-0013802.

PR 14-AUG-1997; 97FR-0010387.

PA (INSR ) INSR ROUSSY GUSTAVE.

PI Bellet D, Bidart JM, Troalen F, Mock P;

DR WPI; 1999-181038/15.

DR N-PSDB; AAX27490.

PS Claim 1; Fig 1; 119pp; French.

XX This sequence represent the early placental insulin-like (EPIL) protein  
 CC encoded by the INSU-4 (insulin-like gene 4) gene. The polypeptide,  
 CC antibodies to the polypeptide, vectors containing the coding sequence and  
 CC probes derived from the coding sequence, can be used to treat tumours,  
 CC preferably angio proliferative tumours, especially Kaposi's sarcoma,  
 CC tumours of the pancreas, liver, uterus or breast, angiosarcomas,  
 CC glioblastomas, neuroblastomas, rhabdomyosarcomas or leiomyosarcomas;  
 CC to promote vascularisation of specific tissues; to treat retinopathy,  
 CC macular degeneration, psoriasis, endometriosis, rheumatoid arthritis,  
 CC atherosclerosis or hypertenoidism; to treat post-angioplasty  
 CC restenosis; to promote or inhibit embryo implantation; to prevent and/or  
 CC treat disorders directly or indirectly connected with insulin-like  
 CC activity; to prevent and/or treat disorders directly or indirectly  
 CC connected with a dysfunction in carbohydrate metabolism, especially  
 CC connected with hypo glycaemia or hyperglycaemia, especially gestational  
 CC diabetes and diabetic complications, especially cardiovascular  
 CC complications.

XX SQ Sequence 139 AA;

XX Alignment Scores:

Pred. No.: 3,856-68 Length: 139  
 Score: 619.00 Matches: 114  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.48% Indels: 0  
 DB: 20 Gaps: 0

US-09-518-842-1\_COPY\_76\_417 (1-342) x AAW99574 (1-139)

QY 1 GCAGAGCTGAGGAGATGTGTCGCCGATTTGGAACAACCTTGTCATATTGCCCATG 60  
 |||||||

Db 26 AlagIleuArGlyCysGlyProArGpHeGlyLysHISleuLeuSerTYrCysPromet 45  
 QY 61 CCTGAGAGACATTTACACACACCCAGAGCGGTGCTGCGAATCTGACGTCCCAA 120  
 |||||||  
 Db 46 ProGluLysThrPheThrThrProGlyIleuLeuGluSerGlyArgProLys 65  
 |||||||  
 QY 121 GAATGCTGTCAACCTCCACACACAGAGACAGCGCTTACAGTACATCGAATTC 180  
 |||||||  
 Db 66 GluMetValSerThrSerAsnAsnLysAspGlyAlaLeuGlyThrThrSerGluPhe 85  
 |||||||  
 QY 181 ATTCCTAATTTGTACACGAGCTGAGAAACCACTGTCTGAGAGGACCATCATTTGAG 240  
 |||||||  
 Db 86 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyInProSerLeuLys 105  
 |||||||  
 QY 241 AAATTAATACCTTCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 |||||||  
 Db 106 LysIleIleLeuSerArgLysLysArgSerIleYargHisArgPheAspPropheCys 125  
 |||||||

RESULT 7

ID AAW69170 standard; Protein; 159 AA.

XX AAW69170;

DT 07-OCT-1998 (first entry)

DE N-terminally tagged Zins1 protein.

KW Zins1; human; placenta; placental; pancreatic islet cell proliferation;  
 KM insulin secretion; diabetes; therapy.

OS Homo sapiens.

PN WO9827210-A1.

PD 25-JUN-1998.

PF 16-DEC-1997; 97WO-US23326.

PR 16-DEC-1996; 96US-0033003.

PA (ZYMO ) ZYMOGENETICS INC.

PI Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;

PI Sprungel KH;

DR WPI; 1998-362779/31.

DR N-PSDB; AAV44664.

PT New isolated protein, Zins1 - obtained from human placenta, which  
 PT increases the proliferation of pancreatic islet cells, used for  
 PT treating diabetes

XX Example 1; Page 62-63; 77pp; English.

XX This sequence is a N-terminally tagged version of the human Zins1  
 CC protein of the invention. The Zins1 protein was isolated from human  
 CC placenta, and is believed to be a new version of the mature protein of  
 CC placenta, having disulphide bonded A and B chains. The protein can be  
 CC used for stimulating the proliferation of pancreatic islets to increase  
 CC insulin secretory capacity of mammals. In particular it can be used for  
 CC the treatment of diabetes. It can also be used for stimulating in vitro  
 CC proliferation of pancreatic islet cells. It can also be used for  
 CC production of antibodies and in detection and diagnosis.

XX SQ Sequence 159 AA;

XX Alignment Scores:

Pred. No.: 4,056-68 Length: 159



XX Hudson PJ, Haley JD, Niall HD, Shine J;  
 XX WPI: 1983-748587/35.  
 DR N-PSDB; AAN30196.  
 XX  
 PT Genes and DNA transfer vectors for prorelaxin expression - useful  
 PT in produ. of porcine relaxin for veterinary and human use  
 XX  
 XX Disclosure; Fig 5; 50pp; English.  
 XX  
 CC The inventors claim synthetic porcine preprorelaxin and prorelaxin  
 CC and synthetic A, B and C peptide chains of prorelaxin, and a gene for  
 CC expression of porcine preprorelaxin or prorelaxin, and their sub-  
 CC units (see AAN30186). They also claim a double-stranded DNA fragment  
 CC for the expression of the signal peptide chain of porcine  
 CC preprorelaxin comprising a coding strand and a complementary strand  
 CC corresp. to a defined mRNA sequence (see AAN30187-N30194) which  
 CC corresp. to the most homologous regions between the pig and rat cDNA  
 CC sequences. A probe (AAN30195) is also claimed.  
 XX  
 SQ Sequence 182 AA:

Alignment Scores:  
 Pred. No.: 1,34e-08 Length: 182  
 Score: 141.00 Matches: 41  
 Percent Similarity: 40.00% Conservative: 7  
 Best Local Similarity: 34.17% Mismatches: 32  
 Query Match: 22.20% Indels: 40  
 Gaps: 2

US-09-518-842-1\_COPY\_76\_417 (1-342) x AAP30392 (1-182)

QY - 100 CTGGAATCTGAGCTCCCAAGAAATGCTGACCTCCACACAAAGATGGACAGCC 159  
 |||||:::||||| ||| ::| :||| ||||| ::|  
 Db 63 LeuGIuThcIyProProAlaGIuThrMetProSerSerIleThrLysAspAlaGIu 82  
 QY 160 TTAGTAGACATCAGCATTCATTCCTAATTGTCACACAGAGCTGAAGAACCTGCT 219  
 ||| |||||:::||||| ||| ||||| ||||| |||||  
 Db 83 LeuLysMetMetLeuGIuThcValProAsnLeuProGIuLLeuLysAlaThrLeuSer 102  
 QY 220 GAAGGCGACGACATTCATTCG----- 237  
 ||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 103 GIuArgGIuProSerLeuArgGIuLeuGIuInSerAlaSerLysAspSerAsnLeuAsn 122  
 QY 238 -----AAGAAATATATACTT----- 252  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 123 PheGIuGIuPheLysLysIleIleLeuAsnArgGIuAsnGIuAlaGIuAspLysSerLeu 142  
 QY 253 -----TCCCGCAAAAGAGAGTGACGTCAC 279  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 143 LeuGIuLeuLysAsnLeuGIuLysAspLysHisSerArgLysArgLeuPheArgMet 162  
 QY 280 AGATTGATTCATTCGTTGAGAGTAATTGTGACGAGATGACGAACTTCAGTTAATTATGT 339  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 163 ThrLeuSerGIuLysCysCysGlnValGIuLysIleArgLysAspIleAlaArgLeuGly 182

RESULT 10  
 AAR31958  
 ID AAR31958 standard; Protein; 178 AA.  
 XX  
 AC AAR31958;  
 XX  
 DT 17-DEC-2001 (updated)  
 DT 09-JUN-1993 (first entry)  
 XX  
 DE Squamous cell specific SQ10 protein.  
 XX  
 KM Squamous cell carcinoma; SCC; marker; differentiation;  
 KW preprorelaxin; rabbit.  
 XX  
 OS Oryctolagus cuniculus.  
 XX

PN USN7783046-N.  
 XX  
 PD 15-DEC-1992.  
 XX  
 PF 28-OCT-1991; 91US-0783046.  
 XX  
 PR 28-OCT-1991; 91US-0783046.  
 XX  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.  
 XX  
 PI Jetten AM, Lotan R;  
 XX  
 DR WPI: 1993-067425/08.  
 DR N-PSDB; AAQ36775.  
 XX  
 PT Diagnosis and assessment of squamous cell carcinoma - by  
 PT identification of intracellular and secreted molecular markers of  
 PT squamous cell differentiation  
 XX  
 PS Example 1; Fig 1; 38pp; English.  
 XX

A cDNA library was constructed from poly(A)+ RNA isolated from  
 CC squamous differentiated rabbit tracheal epithelial cells. The cDNA  
 CC clone, SQ10, isolated from this library had a sequence homologous  
 CC with those of the genes encoding human and porcine relaxin, i.e. the  
 CC gene encodes a secreted protein related to preprorelaxin. The  
 CC protein is a marker for squamous cell differentiation, and is  
 CC secreted extracellularly. This allows body fluids e.g. sera, saliva  
 CC and urine to be analysed in diagnosis of squamous cell carcinoma (SCC).  
 CC Antibodies raised against the protein or immunogenic fragments  
 CC react with squamous differentiated cells but not undifferentiated cells  
 CC and are thus specific for a marker of squamous cell differentiation.  
 CC See also AAR31958.

(NOTE: Revised entry submitted to correct the patent number format of  
 CC US Government-owned NTIS applications to prevent clashes with ongoing US  
 CC granted patent numbers. For further information please visit the Derwent  
 CC web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)

SQ Sequence 178 AA:

Alignment Scores:  
 Pred. No.: 7,22e-05 Length: 178  
 Score: 111.00 Matches: 39  
 Percent Similarity: 32.81% Conservative: 3  
 Best Local Similarity: 30.47% Mismatches: 42  
 Query Match: 17.48% Indels: 44  
 Gaps: 3

US-09-518-842-1\_COPY\_76\_417 (1-342) x AAR31958 (1-178)

QY 37 CACTTGCTGTCATTAATGCGCCATGCGAGAGACATTCACACCACCCAGAGGCTGG 96  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 50 HisLeuGIuArgGIuSerProSerProGIuAsnProHe----- 62  
 QY 97 CTGCTGAATCTGAGCGCTCCCAAGAAATGCTGTCACCTCCCAACAAGATGGACAA 156  
 ||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 63 ---LeuSerSerGIuProAlaIaGIuThrValProSerSerIleLysLysAspAlaGIu 81  
 QY 157 GCCTTAGTAGTACATCAGATTAATTCCTAATTGTCACACAGAGCTGAAGAACCACTG 216  
 ||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 82 AsnAlaAsnThrMetLeuGIuSerIleProAsnLeuProGIuLLeuThrAlaThrLeu 101  
 QY 217 TCTGAAGGCGACGACATCA----- 234  
 ||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 102 PheGIuLysGIuProSerLysLeuTYrLeuGIuInTYrLeuProThrLeuLysLysSerAsn 121  
 QY 235 -----TTGAAGAAATATA----- 249  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 122 ValSerPheGIuGIuPheLysLysIleIleGIuAsnIleGIuArgGIuValGIuGIuLys 141  
 QY 250 -----CTTCCCGCAAAAGAGAGTGACGCTCAGACATTTGAT 288  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 142 SerAlaSerGIuSerAsnThrPheSerArgLysLysArgGIuPheSerGIuSerLeuPro 161

QY 289 CCATTCTGCTGAAGTAATTGT 312  
 DB 162 GUGlucCysCysLysTYRGLYCys 169  
 RESULT 11  
 AAR64904  
 ID AAR64904 standard; Protein: 150 AA.  
 AC AAR64904;  
 XX  
 DT 06-SEP-1995 (first entry)  
 DE Prorelaxin fragment amino acids 12-161.  
 XX  
 XX Prorelaxin; cervical ripening; ovarian peptide hormone;  
 KW mammary gland development; skin elasticity; cardiovascular therapy;  
 KW relaxin; premature labour; plasmid pTR21.  
 OS Synthetic.  
 XX  
 PN WO9500645-A.  
 XX  
 PD 05-JAN-1995.  
 XX  
 PF 20-JUN-1994; 94MO-US06997.  
 XX  
 PR 21-JUN-1993; 93US-0080354.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Breese T, Hayenga K, Rinderknecht E, Vandlen R;  
 PI Yansura D;  
 DR WPI: 1995-052082/07.  
 DR N-PSDB: AAO76309.  
 XX  
 PT Relaxin prodn. from non-natural recombinant prorelaxin - by  
 PT cleavage of leader and C peptide chain, also new prorelaxin and  
 PT related DNA, vectors etc.  
 XX  
 PS Example 1; Fig 9a; 62pp; English.  
 XX  
 CC AAO76309 is the Not I-Bam HI restriction fragment of the plasmid  
 CC pTR21, which encodes AAR64904, prorelaxin amino acids 12-161.  
 CC Relaxin is produced by removing the non-naturally occurring leader  
 CC and C-peptide from PR with a cleavage agent. Relaxin is an ovarian  
 CC peptide hormone involved in the inhibition of premature labour,  
 CC cervical ripening and the development of the mammary glands. It  
 CC may also improve skin elasticity and has been used in cardiovascular  
 CC therapy.  
 CC  
 SQ Sequence 150 AA:  
 Alignment Scores:  
 Pred. No.: 0.000328 Length: 150  
 Score: 105.50 Matches: 43  
 Percent Similarity: 35.67% Conservative: 13  
 Best Local Similarity: 27.39% Mismatches: 44  
 Query Match: 16.61% Indels: 57  
 DB: 16 Gaps: 6  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x AAR64904 (1-150)  
 QY 1 GCAGAGCTGAGGGATGTGT---CCCCATTGGAAAAACCTTCGTCATATTGCCCC 57  
 DB 7 AAGlInIleAlaIleGlySerThrTrpSerLysArgSerLeuSerLInLasp 26  
 QY 58 ATGCCGTGAAGACATTCACACACACCCAGAGGGGTGGCTGTGGAATCTGACGTCC 117  
 DB 27 AlapProGln-----ThrPro-----ArgPro 33  
 QY 118 ---AAAGAAATGCTGTCACCTCCAAACAAGATGACAAAGCTTAAAGTAGACATCA 174

DB 34 ValAlaGlnIleValProSerPheIleAsnLysAspTrpGluThrIleAsnMetSer 53  
 QY 175 GAATTCATTCCTAATTGTCCACAGAGCTGAAGAAACACATCTCTGAAAGGCACATCA 234  
 DB 54 GluPheValAlaAsnLeuProGlnLeuLysLeuThrLeuSerGluMetGlnProAla 73  
 QY 235 TTG-----  
 DB 74 LeuProGlnLeuGlnGlnHisValProValLeuLysAspSerSerLeuPheGluGlu 93  
 QY 238 ---ACGAAATTAATCTT-----  
 DB 94 PheLysLysLeuIleArgAsnArgInserGluAlaAlaAspSerSerProSerGluLeu 113  
 QY 253 -----TCCCGCAAAAGAGAGAGTGGACGTCACAGATTGAT 288  
 DB 114 LysTYRLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTYRSerAlaLeuAla 133  
 QY 289 CCATTCTGCTGAAGTAATTGTGACGATGACACTTCGTAATTATGT 339  
 DB 134 AsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 150  
 RESULT 12  
 AAR07987  
 ID AAR07987 standard; protein: 162 AA.  
 AC AAR07987;  
 XX  
 DT 27-FEB-1991 (first entry)  
 DE H2 prorelaxin gene encoded by insert in plasmid pTRProRelasp.  
 XX  
 KW Relaxin; PR.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9013659-A.  
 XX  
 PD 15-NOV-1990.  
 XX  
 PF 16-APR-1990; 90MO-US02085.  
 XX  
 PR 04-MAY-1989; 89US-0347550.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Hennen DJ, Vandlen RL, Wikins JA, Yansura D;  
 PI N-PSDB: AAO06595.  
 DR  
 DR  
 PT Cleaving polypeptide into polypeptide cleavage prods. - by  
 PT treating free-cysteine form of polypeptide at desired junction.  
 XX  
 PS Disclosure; Fig 2a; 56pp; English.  
 XX  
 CC The plasmid was prepd. from a clone isolated from a cDNA library  
 CC prepd. from RNA isolated from human corpus luteum, screened with an  
 CC HI-cDNA probe. A fragment encoding the N-terminal of PR was iso-  
 CC lated and was ligated with a 410 bp fragment encoding AAs 17-153 of  
 CC PR, and a fragment from pHCN2071\*1L (including the Trp promoter, AAs  
 CC 1-137 of methGH, and amp and tet resistance. A portion of the  
 CC resulting construction (pPEPRH2) was ligated with a fragment from  
 CC the original clone encoding the PR C-terminal, and a fragment from  
 CC pBR322XAP encoding part of the beta-lactamase gene. This produced  
 CC plasmid pTRProRel. A Still signal sequence was then fused to the  
 CC PR gene and a portion including the gene and signal sequence was  
 CC ligated into a vector identical to pTRPSIIGH in which the HGH  
 CC gene had been removed. The resulting construction, pTRPSIIPROEL  
 CC was treated to remove the Still sequence and the first 11 AAs of H2  
 CC PR. This was replaced with a synthetic DNA duplex encoding the  
 CC first 12 AAs of H2 PR (including Asp1), to produce pTRProRelasp.

CC the coding part of which is shown below. The vector can be used to  
 CC construct vectors that encode Asp-inserted H2PR with enhanced acid  
 CC cleavage sites.  
 CC See also AAR07988.

SO Sequence 162 AA;

Alignment Scores:

Pred. No.: 0.000337 Length: 162  
 Score: 105.50 Matches: 43  
 Percent Similarity: 35.67% Conservative: 13  
 Best Local Similarity: 27.39% Mismatches: 44  
 Query Match: 16.61% Indels: 57  
 DB: 11 Gaps: 6

US-09-518-842-1\_COPY\_76\_417 (1-342) x AAR07987 (1-162)

QY 1 GCAGAGCTGAGGGAGTGTGTTCCCGATTGGAAAACACTTGTCTCATATTGCCCC 57  
 DB 19 AlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGluAsp 38  
 QY 58 ATGCCTGAGAGACATTAC 117  
 DB 39 AlaProGln-----ThrPro-----ArgPro 45  
 QY 118 ---AAAGAAATGGTCAACCTCCACACAAACAAAGATGACAGCCTTAGTACGACATCA 174  
 DB 46 ValAlaGlnIleValIleProSerPheIleAsnLysAspThrGluThrIleAsnMetMetSer 65  
 QY 175 GAATTCATTCCTTAATTTGTCCACAGAGCTGAGAGAAACACACTGTCTGAAGGCGACCATCA 234  
 DB 66 GluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluMetGlnProAla 85  
 QY 235 TTG----- 237  
 DB 86 LeuProGlnLeuGlnGlnHisValIleProValLeuLysAspSerSerLeuLeuPheGlnGlu 105  
 QY 238 ---AAGAAATATTAATCT----- 252  
 DB 106 PheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerSerProSerGluLeu 125  
 QY 253 -----TCCCGCAAAAAGAGAGTGGACGTCACACATTTGAT 288  
 DB 126 LysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaLeuAla 145  
 QY 289 CCATTCCTGTGGAAGTAATTTGTGACGATGGAACCTTCAGTTAATTATGT 339  
 DB 146 AsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 162

RESULT 13  
 AAP94621  
 ID AAP94621 standard; protein, 185 AA.  
 XX  
 AC AAP94621;  
 XX  
 DT 21-JUN-1990 (first entry)  
 XX  
 DE Amino acid sequence of human preprorelaxin H2.  
 XX  
 KM Relaxin; H2-relaxin; public symphys.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..25 /label=Signal peptide  
 FT 26..57 /label=B-chain  
 FT 58..161 /label=C-peptide  
 FT 162..159 /label=A-chain  
 FT /label=A-chain  
 XX

PN EP303033-A.  
 XX  
 PD 15-FEB-1989.  
 XX  
 PF 12-DEC-1983; 83EP-0110103.  
 XX  
 PR 13-DEC-1982; 82AU-0007247.  
 XX  
 PA (FLOR-) FLORE H INST EXPER.  
 XX  
 PI Hudson PJ, Tryear GW, Miall HD;  
 XX  
 DR WPI; 1989-047874/07.  
 XX  
 DR N-PSDB; AAN92483.  
 XX  
 PT New human H2-relaxin analogues -  
 PT with shortened and/or modified A and/or B chains.  
 CC  
 CC Disclosure; 25pp; English.  
 CC  
 CC "H2" relaxin is deduced from a cDNA clone, has the general properties of  
 CC a growth factor and is capable of altering nature of connective tissue  
 CC and inducing smooth muscle contraction, specifically during labour.

SQ Sequence 185 AA;

Alignment Scores:

Pred. No.: 0.000354 Length: 185  
 Score: 105.50 Matches: 43  
 Percent Similarity: 35.67% Conservative: 13  
 Best Local Similarity: 27.39% Mismatches: 44  
 Query Match: 16.61% Indels: 57  
 DB: 10 Gaps: 6

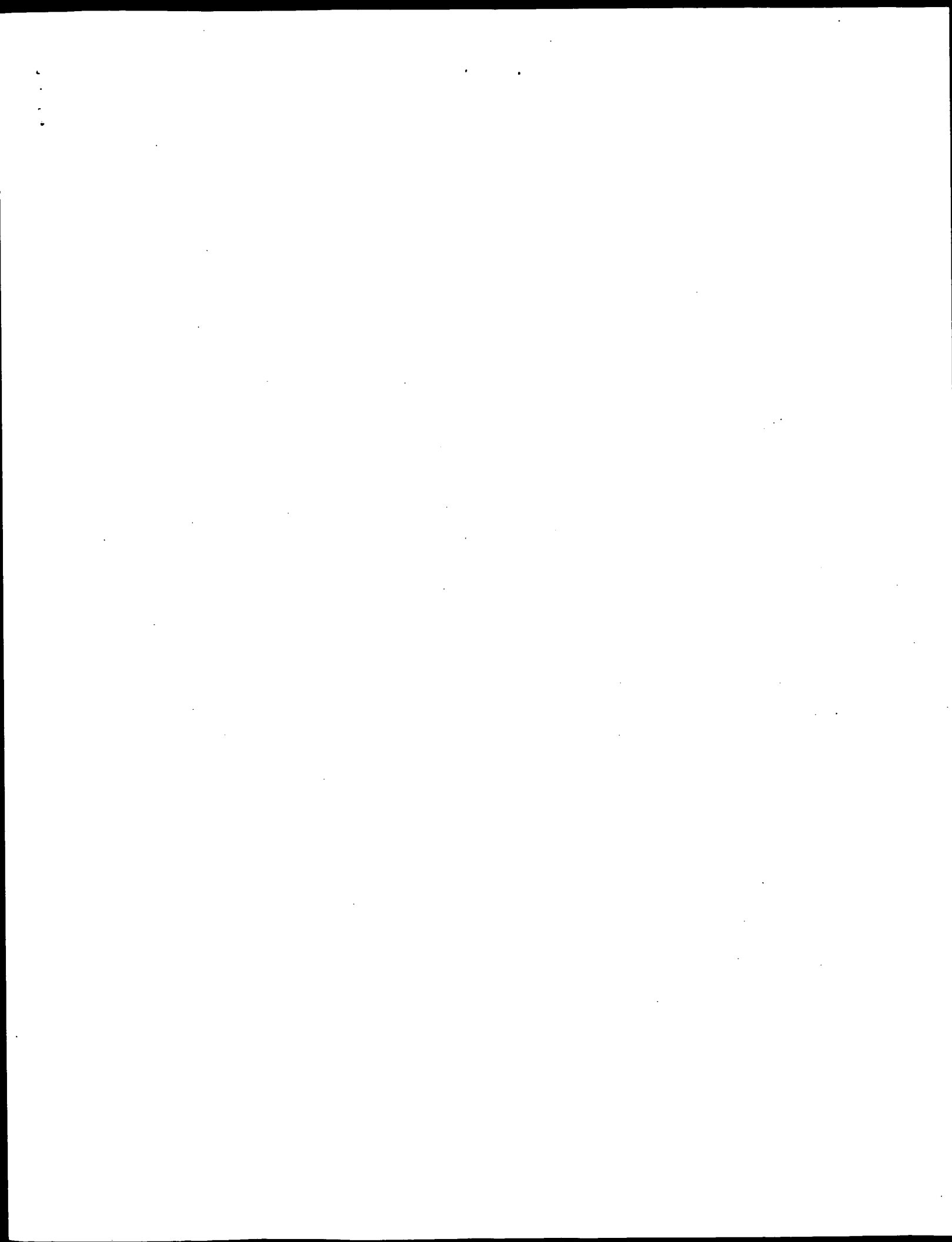
US-09-518-842-1\_COPY\_76\_417 (1-342) x AAP94621 (1-185)

QY 1 GCAGAGCTGAGGGAGTGTGTTCCCGATTGGAAAACACTTGTCTCATATTGCCCC 57  
 DB 42 AlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGluAsp 61  
 QY 58 ATGCCTGAGAGACATTAC 117  
 DB 62 AlaProGln-----ThrPro-----ArgPro 68  
 QY 118 ---AAAGAAATGGTCAACCTCCACACAAACAAAGATGACAGCCTTAGTACGACATCA 174  
 DB 69 ValAlaGlnIleValIleProSerPheIleAsnLysAspThrGluThrIleAsnMetMetSer 88  
 QY 175 GAATTCATTCCTTAATTTGTCCACAGAGCTGAGAGAAACACACTGTCTGAAGGCGACCATCA 234  
 DB 89 GluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluMetGlnProAla 108  
 QY 235 TTG----- 237  
 DB 109 LeuProGlnLeuGlnGlnHisValIleProValLeuLysAspSerSerLeuLeuPheGlnGlu 128  
 QY 238 ---AAGAAATATTAATCT----- 252  
 DB 129 PheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerSerProSerGluLeu 148  
 QY 253 -----TCCCGCAAAAAGAGAGTGGACGTCACACATTTGAT 288  
 DB 149 LysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaLeuAla 168  
 QY 289 CCATTCCTGTGGAAGTAATTTGTGACGATGGAACCTTCAGTTAATTATGT 339  
 DB 169 AsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 185

RESULT 14  
 AAP40154  
 ID AAP40154 standard; protein, 185 AA.  
 XX  
 AC AAP40154;

[illegible][illegible]







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 14, 2003, 19:37:50 ; Search time 19 Seconds  
(without alignments)  
3460.839 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 635  
Sequence: 1 GCAGAGCTGAGGGGATGTGG.....CTTCAGTAATTAATGATACA 342

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=/c9n2\_1/USPTO.spool/US0518442/rnat\_14062003.175811.12284/app.query.fasta\_1.519  
-DB=PIR\_73 -OPMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -IOOPT=0 -IOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cbi -LIST=45  
-DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US0518442.c9cn\_1.1.62.etrnat\_14062003.175811.12284 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	22.2	182	1	relaxin precursor
2	116.5	18.3	185	2	relaxin precursor
3	111	17.5	178	2	20K protein - rbb
4	105.5	16.6	166	2	relaxin 2 precursor
5	105.5	16.6	185	1	relaxin B,C and A
6	103.5	16.3	143	2	relaxin 1 precursor
7	103.5	16.3	185	1	relaxin 1 precursor
8	103.5	16.3	185	2	relaxin 1 precursor
9	101	15.9	560	2	relaxin 1 precursor
10	100.5	15.8	166	2	relaxin 1 precursor
11	97.5	15.4	186	1	relaxin 1 precursor
12	83	13.1	321	2	relaxin precursor
13	78	13.0	754	2	hypothetical prote
14	77	12.1	457	2	probable competent
					protease (EC 3.4

15	76.5	12.0	160	2	relaxin - guinea p
16	74.5	11.7	1531	2	probable outer mem
17	73.5	11.6	273	2	NAD synthetase, pr
18	73.5	12.2	402	2	probable disease r
19	73	12.2	993	2	26S proteasome reg
20	73	11.5	1435	2	regulatory protein
21	72.5	11.4	292	2	hypothetical prote
22	72.5	11.4	1259	2	hypothetical prote
23	72.5	12.1	1744	1	complement C4a pre
24	72	11.3	391	2	replication acid rece
25	72	11.3	453	2	retinoic acid rece
26	72	11.3	483	2	succinate-semialde
27	72	11.3	752	2	hypothetical prote
28	72	11.3	836	1	DNA-directed DNA p
29	72	11.3	2397	2	verstein precursor
30	71.5	11.9	1217	2	hypothetical prote
31	71	11.8	206	2	hypothetical prote
32	71	11.2	1490	2	nonstructural poly
33	70.5	11.1	271	1	gene 18 protein -
34	70.5	11.1	347	2	hypothetical prote
35	70.5	11.8	766	2	hypothetical prote
36	70.5	11.1	957	2	excinuclease ABC c
37	70	11.0	283	2	probable pantocate-
38	69.5	10.9	479	2	GTP1/Obg-family GT
39	69.5	10.9	645	2	threonyl-tRNA synt
40	69.5	10.9	1571	2	lactobacillus phag
41	69	11.5	274	1	nonstructural prot
42	69	11.5	349	2	hypothetical prote
43	69	10.9	508	2	probable long chai
44	69	11.5	544	2	probable substrate
45	69	11.5	570	2	probable kexin (EC

ALIGNMENTS

RESULT 1  
XAPG  
relaxin precursor - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 18-Jun-1999  
C:Accession: A90934; A93187; A90205; A90196; S32313; S32312; A29796; A01615  
R:Halley, J.; Hudson, P.; Scallan, D.; John, M.; Shine, J.; Tregear, G.; NI  
DNA 1, 155-162, 1982  
A:Title: Porcine relaxin: molecular cloning and cDNA structure.  
A:Reference number: A90934; MUID:83157118; PMID:6897721  
A:Accession: A90934  
A:Molecule type: DNA  
A:Residues: 1-182 <HAL>  
A:CROSS-references: GB:K01088; MID:g164634; PIDN:AAA31114.1; PID:g164635  
R:James, R.; Nall, H.; Kwok, S.; Bryant-Greenwood, G.  
Nature 267, 544-546, 1977  
A:Title: Primary structure of porcine relaxin: homology with insulin and related grow  
A:Reference number: A93187; MUID:77213067; PMID:876374  
A:Accession: A93187  
A:Molecule type: Protein  
A:Residues: 25-50, TWGR; 161-182 <TAM>  
R:Schwabe, C.; McDonald, J.K.; Steinetz, B.G.  
Biochem. Biophys. Res. Commun. 75, 503-510, 1977  
A:Title: Primary structure of the B-chain of porcine relaxin.  
A:Reference number: A90205; MUID:77157271; PMID:851452  
A:Accession: A90205  
A:Molecule type: protein  
A:Residues: 25-47, VW; 50 <SCH>  
R:Schwabe, C.; McDonald, J.K.  
Biochem. Biophys. Res. Commun. 74, 1501-1504, 1977  
A:Title: Demonstration of a pyroglutaminyl residue at the N terminus of the B-chain of  
A:Reference number: A90201; MUID:77141136; PMID:843375  
A:Contents: annotation: pyroglutamate carboxylic acid  
R:Schwabe, C.; McDonald, J.K.; Steinetz, B.G.  
Biochem. Biophys. Res. Commun. 70, 397-405, 1976  
A:Title: Primary structure of the A chain of porcine relaxin.  
A:Reference number: A90196; MUID:76231539; PMID:938497  
A:Accession: A90196

```

A:Molecule type: Protein
A:Residues: 161-169, 'E', 171-182 <SCH>
R:Schwabe, C.; McDonald, J.K.
Science 197, 914-915, 1977
A:Title: Relaxin: a disulfide homolog of insulin
A:Reference number: A94245; MUID:77236040; PMID:887933
A:Contents: annotation: disulfide bonds
R:Kohsaka, T.; Takahara, H.; Sugawara, K.; Tagami, S.
Biol. Chem. Hoppe-Seyler 374, 203-210, 1993
A:Title: Endogenous heterogeneity of relaxin and sequence of the major form in pregnant
A:Reference number: S32312; MUID:93257096; PMID:8489740
A:Accession: S32313
A:Molecule type: Protein
A:Residues: 25-53 <KOH>
A:Accession: S32312
A:Molecule type: protein
A:Residues: 161-182 <KO2>
R:Haley, J.; Crawford, R.; Hudson, P.; Scanlon, D.; Tregear, G.; Shine, J.; Niall, H.
J. Biol. Chem. 262, 11940-11946, 1987
A:Title: Porcine relaxin. Gene structure and expression.
A:Reference number: A29796; MUID:87308187; PMID:2442155
A:Accession: A29796
A:Molecule type: DNA
A:Residues: 1-115, 'L', 117-182 <HA2>
A:Cross-references: GB:J02792; NID:9164636; PIDN:AAA31115.1; PID:9164637
C:Comment: Relaxin is an ovarian hormone that acts with estrogen to produce dilation of
C:Genetics:
A:introns: 70/1
C:Superfamily: Insulin
C:Keywords: pyroglutamic acid
F:1-24/Domaln: signal sequence #status predicted <SIG>
F:25-56/Domaln: relaxin chain B #status experimental <RCB>
F:161-182/Domaln: relaxin chain A #status experimental <RCA>
F:25-56/161-182/Product: relaxin #status experimental <MAT>
F:161-182/Domaln: relaxin chain A #status experimental <RCA>
F:25/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experime
F:34-169,46-182,168-173/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.: 2, 74e-07 Length: 182
Score: 141.00 Matches: 41
Percent Similarity: 40.00% Conservative: 7
Best Local Similarity: 34.17% Mismatches: 32
Query Match: 22.20% Indels: 40
DB: 1 Gaps: 2

US-09-518-842-1_COPY_76_417 (1-342) x RXPG (1-182)

QY 100 CTGGAATCTGGACGTCCTCCAAAGAAATGGTGTCACTCCAAACAAAGATGACAGCC 159
|||||:||||| ||| ||| ::| :||| ||||| :|||
Db 63 LeuGIuThrGlyProProlaGluThrMetProSerSerIleThrIysaspAlaGluIle 82
160 TTAGAGCAGACATCGAATTCATTCCTATTATTGTCCAGCAGCTGCAAGAACCATGTCT 219
|||||:||||| |||||:||||| |||||:||||| |||||
Db 83 LeuIysMetMetLeuGluIuPheValIProAsnLeuProGlnGluLeuIysAlaThrIreLeuSer 102
220 GAAGGCACACCATCATTTG----- 237
||| |||||:||||| |||
Db 103 GluArgGlnProSerIreLeuArgGluLeuGlnIleAsnIleAlaSerIysaspSerAsnLeuAsn 122
238 -----AAGAAATTAATACTT----- 252
|||||:||||| |||||:||||| |||||
Db 123 PheGluGluPheIuPheIysIleIleLeuAsnAlaGlnAsnGluAlaGluIuAspIysSerLeu 142
253 -----TCCCGCAAAAAGAGAGGTGACGCTCAC 279
|||||:||||| |||||:||||| |||||
Db 143 LeuGluLeuIysAsnIleuGlyLeuAspIysSHISerArgIysIysArgIeuPheArgMet 162
280 AGATTGATTCATCTTCTGTGTGAGAGTAATTGTGACAGATGGAACCTCAGTTAATATATCT 339
|||||:||||| ||| |||||:||||| |||
Db 163 ThrLeuSerGluIuIysCysGlnValGlyCysIleArgIysAspIleAlaArgLeuCys 182

```

relaxin precursor - mouse

C:Species: Mus musculus (house mouse)  
C>Date: 07-May-1995 #sequence revision 21-Jul-1995 #text\_change 16-Jul-1999  
C:Accession: S48082; PC2067; PNM0626  
R:Evans, B.A.; John, M.; Fowler, K.J.; Summers, R.J.; Cronk, M.; Shine, J.; Tregear,  
J. Mol. Endocrinol. 10, 15-23, 1993  
A:title: The mouse relaxin gene: nucleotide sequence and expression.  
A:Reference number: S48082; MUID:93199663; PMID:8452637  
A:Accession: S48082  
A:Molecule type: mRNA  
A:Status: preliminary  
A:Residues: 1185 <EVA>  
A:Cross-references: EMBL:Z27086; NID:g414780; PIDN:CA81611.1; PID:g414781  
R:Buellesbach, E.E.; Schwabe, C.  
Biochem. Biophys. Res. Commun. 196, 311-319, 1993  
A:title: Mouse relaxin: synthesis and biological activity of the first relaxin with a  
A:Reference number: PNM0626; MUID:94030011; PMID:8216305  
A:Accession: PC2067  
A:Molecule type: protein  
A:Residues: 23-57 <BUO>  
A:Accession: PNM0626  
A:Molecule type: protein  
A:Residues: 161-185 <BU2>  
A>Note: proteins with and without 184-Tyr were synthesized, their biological activity  
C:keywords: hormone  
F:23-57,161-185/Product: relaxin #status experimental <MAT>  
F:23-57/Domain: chain B #status experimental <CHB>  
F:161-185/Domain: chain A #status experimental <CHA>  
F:36-171,48-185,170-175/Dissulfide bonds: #status experimental

Alignment Scores:

Pred. No.:	0.000147	Length:	185
Score:	116.50	Matches:	38
Percent Similarity:	35.76%	Conservative:	16
Best Local Similarity:	25.17%	Mismatches:	56
Query Match:	18.35%	Indels:	41
DB:	2	Gaps:	2

US-09-518-842-1\_COPY\_76\_417 (1-342) x S48082 (1-185)

```
OY      7 CTGAGGGATGTGTCGCCCATTTGGAAACACTTGCCTCATATTGGCCCATGCCTGAG 66
           ::::| | | | | | | | : : : : : | | | | |
Db     33 lIeaYMeCySgIyArGIuTrtAlaRGluLeuIlElySlICySGlAlASerVal 52
OY     67 AAGACATTCACCACACCCAGGAGCGTGCGTGAATCGACGTCGCCAAGAANTG 126
           :: : : : : : : : : : : : : : : : : : : : :
Db     53 GlAyArgLeuAlaleuSerGlnGlIngluPrAlaLeuAlaArgGlnAlAtHcgluVal 72
OY    127 GTGTCACCTCCACACAACAAGANTGACACGCTTAGTACGACATTCAGATTCTCT 186
           ||| : : : | | | | | | | : : : | | | | | : : : | |
Db     73 ValProSerPheleAsnIlySnPsAlaGluPrOpheaSPrtThrLeuIlyCSylEuRo 92
OY    187 AATTGTACCAGCATGTAAGAACCACTGCTGTGAAGGCGACGCTCATTAAGAAMAATA 246
           ||||| | | | | | | | | | | | | | | | | | | | | |
Db     93 AsnIeuSerGluInleuIlySnAlValleuSerGluAlaSerLeuProGluLeu 112
OY    247 -----ATPACTT----- 252
           : : : | | |
Db    113 GlnHisAlaProValleuSerAspSerValIalSerLeuGluGlyPheIlylsYThrLeu 132
OY    252 ----- 252
Db    133 HisAspArgLeuGlyGlnAlaGluAspGlySerProProGlyLeuIlyStYLeuInSer 152
OY    253 -----TCCCAGAAAAGAGAAGTGCACGCTCACAGATTTGATTCATCTGTTGTGAA 303
           ||||| | | | | | | | | | | | | | | | | | | | | |
Db    153 AspTrnHisSerArgIlysnArgIlysnSerGlyGlyLeuMetSerGlnIncIynScyShIs 172
OY    304 GPAAATTGTGACGATGACGACTTCAGTTAAATA 336
           ||| | | | | | : : : | | | | | |
Db    173 ValGlyCySerArgArgSerIleaIalIyLeu 183
```

RESULT 3  
A49014  
20K protein - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Jul-1999  
C:Accession: A49014  
R:Jetterton, A.W.; Bennacki, S.H.; Floyd, E.E.; Saunders, N.A.; Pieniazek, J.; Lotan, R.  
Cell Growth Differ. 3, 549-556, 1992  
A:Title: Expression of a p190ret/axlin-like gene during squamous differentiation of rabbit  
A:Accession number: A49014; MUID:93002619; PMID:1339318  
A:Accession: A49014  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-178 <JER>  
A:Cross-references: GB:54594.0; NID:q257388; PIDN:AAB23548.1; PID:q257389  
A:Experimental source: tracheobronchial epithelial cells  
A:Note: sequence extracted from NCBI Backbone (NCBIN:115816, NCBIPI:115821)  
C:Superfamily: Insulin

Alignment Scores:	
Pred. NO.:	0.0006
Score:	111.00
Percent Similarity:	32.81%
Best Local Similarity:	30.47%
Query Match:	17.48%
DB:	2
	Gaps:
	3

```

QY 37 CACTTGCCTGCATATTATGCCCCCATGCTGTAGAGAGACATTCCACACCACCCACGAGAGGGTGG 96
Db 50 HistLeuGluArgIuSerProSerProGluAsnProPhe-----62
QY 97 CTCCTGGAATCTTGAGAGCTCCCAAAGAAATGGTGTCAACCTCCACMCACAAAGATGACAA 156
Db 63 ---LeuSerSerGlyProAlaIaGluThrValProSerSerIleLysAlaAspAlaGlu 81
QY 157 GCCTTAGTAGACGACATCAGAAATTCATTCCTATTGTTGTACCCAGAGCTGAGAAACCACTG 216
Db 82 AsnAlaAsnThrMetLeuGluSerIleProAsnLeuProGlnGluLeuThrAlaThrLeu 101
QY 217 TCTGAAGGGCAGCCATCA-----234
Db 102 PheGluLysGlnProSerLysLeuTyrLeuGlnTyrLeuProThrLeuLysSerAsn 121
QY 235 -----TTGAGAAATATAA-----249
Db 122 ValSerPheGluGluPheLysIleIleGlnAsnIleGlnArgIValGlnGlySer 141
QY 250 -----CTTTCGCCGCAAAAAGAGAAATGGAGCTGCACAGATTGAT 288
Db 142 SerAlaSerGluSerAsnThrPheSerIaTgIysArgIaGlnPheSerGluSerLeuPro 161
QY 289 CCATTCTGTTTGGAAGTAATTTGT 312
Db 162 GluGluCysCysLeuTyrGlyCys 169

```

A:Accession: number: S42770  
A:Accession: S42786  
A:Molecule type: mRNA  
A:Residues: 1-166 <EVA>  
C:Cross-references: EMBL:Z27245; NID:9416109; PIDN:CAA81758.1; PID:9416110  
C:Genetics:  
A:Gene: rlx2  
A:Superfamily: insulin

```
F;1-5/Domain: signal sequence (fragment) #status predicted <SIG>
F;6-166/Product: relaxin 2 #status predicted <MAT>
```

Alignment Scores:	
Pred. No.:	0.00244
Score:	105.50
Percent Similarity:	35.67%
Best local Similarity:	27.39%
Query Match:	16.61%
DB:	2
Length:	166
Matches:	43
Conservative:	14
Mismatches:	43
Indels:	57
Gaps:	6

US-09-518-842-1\_COPY\_76\_417 (1-342) x S42786 (1-166)

QY	1	GCACAGCTGAGGAGATGTGT	----	CCCCGATTGGAAACCACTTCCTTCATATATGCCCC	57
Db	23	AlaGlnIleAlaIleCysGlyLysSerThrTrpSerIysArgSerLeuSerGlnuasp			42
QY	58	ATGCTTGAGAAAGACATTTCACCACCA	CCCCAGGAGGTGGCTGGAAATCTGACGTC	117	
Db	43	AlaProGln	-----	ArgPro	49
QY	118	---AAAGAAATGGTGTCAACCTCGAACACAAAGATGACACAAAGCCTTAGTACGACATCA			174
Db	50	ValAlaGluIleValIProSerPheIleAsnLysAspThrGluThrIleAsnMetMetSer			69
QY	175	GAATTCATTCTCTTAATTTGTGCACGAGAGCTGAAGAACCACTGTCGTGAAGGGCACCATCA			234
Db	70	GluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluMetGlnProIa			89
QY	235	TTG	-----		237
Db	90	LeuProGlnLeuGlnGlnIleTyrValIProValLeuLysAspSerSerLeuLeuPheGluGlu			109
QY	238	---AAGAATAATAACTT	-----		252
Db	110	PheLysLysLeuIleArgAsnArgGlnSerGlnAlaIleAspSerSerProSerGluLeu			129
QY	253	-----TCCCGCAAAAGAGAAAGTGGAGCTGCACAGATTGGAT			288
Db	130	LysTyrIleuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaIleuAla			149
QY	289	CCATTCTGGTGTGAGTAATTTGTGACAGATGAGCACTCAGTTAATATATGT			339
Db	150	AsnLysCysLysHisValIleGlyCysThrLysArgSerLeuAlaArgPheCys			166

RESULT 5  
A60982

relaxin 2 precursor [validated] - human  
N;Alternate names: preprorelaxin 2

Species: *Homo sapiens* (man)

C; Date: 03-Mar-1994 #sequence

C;Accession: A05092; A60982

R;Hudson, P.; John, M.; Crawford

EMBO J. 3, 2333-2339, 1984

A::Title: Relaxin gene express

A:Reference number: A05093; M

Accession Number: A03032; 105003

A; ACCESSION: A05092

A; molecule type: mRNA

A;ResIdues: 1-185 &lt;HUD&gt;

A;Cross-references: GB:X00948

R;Stults, J.T.; Bourell, J.H.

Biomed. Environ. Mass Spectrom.

A:Title: Structural character

A: Reference number: A60982. M

Reference number: A00362; P

A; ACCESSION: A6098Z

A; molecule type: protein

A;residues: 25-53;162-185 <ST

C;Genetics:

A;Gene: GDB:RLN2

A: Cross-references: GDB: 11955

A:Map position: 9nter-9a12

C: Superfamily: incultis

C: Keyword: Success, success, achievement

key words: ovaly; pylogluca

100

F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-53/Domain: relaxin 2 chain B #status experimental <BCH>  
 F:54-152/Domain: relaxin 2 #status experimental <AMT>  
 F:58-157/Domain: relaxin 2 connecting C peptide #status predicted <CPEP>  
 F:162-185/Domain: relaxin 2 chain A #status experimental <ACH>  
 F:35-172/47-185/171-176/Disulfide bonds: #status experimental  
 F:162/Modified site: pyrolysine carboxylic acid (Gln) (in mature form) #status experime

## Alignment Scores:

Pred. No.: 0.00247 Length: 185  
 Score: 105.50 Matches: 43  
 Percent Similarity: 35.67% Conservative: 13  
 Best Local Similarity: 27.39% Mismatches: 44  
 Query Match: 16.61% Indels: 57  
 DB: 1 Gaps: 6

US-09-518-842-1\_COPY\_76\_417 (1-342) x A60982 (1-185)

QY 1 GCAGAGCTGAGGGAGTGTGT---CCCCGATTGGAAACACTGTGTCATATTGCCCC 57  
 |||:||||| |||: ||| |||: |||  
 Db 42 Alaglinlelalellecysglmetserthrtipserlysgserleuarglnluasp 61  
 |||:||||| |||: ||| |||: |||  
 QY 58 ATGCTGAGAGAGCATTCACACCCAGAGGGTGGCTGCTGGAATCTGGACGTCC 117  
 |||:||||| |||: ||| |||: |||  
 Db 62 AlaProgin-----ThrPro-----ArgPro 68  
 |||:||||| |||: ||| |||: |||  
 QY 118 ---AAGCAATGGTGTCAACCTCCACACCAAGATGACACCTTAGGTACACATCA 174  
 |||:||||| |||: ||| |||: |||  
 Db 69 Valalaglinlelvalproserphelileasnllyaspthrciluvhrilleasmetser 88  
 |||:||||| |||: ||| |||: |||  
 QY 175 GAATTCATCTCTAATTGTGCACAGAGCTGAAGAACCACTGCTGAGGCGACCATCA 234  
 |||:||||| |||: ||| |||: |||  
 Db 89 GluphevalalalsanleuProginleuLeuLeuThreusergluwetginProAla 108  
 |||:||||| |||: ||| |||: |||  
 QY 235 TTG----- 237  
 |||  
 Db 109 LeuProginleuInglnInHlValProvalleuLysaspserserleuLeuPhegluIn 128  
 |||  
 QY 238 ---AAGCAATTAATCTT----- 252  
 |||:||||| |||: ||| |||: |||  
 Db 129 PheLysLysLeuLeuArgasnArgGlnSerGluAlaAlaAspSerSerProserGluLeu 148  
 |||:||||| |||: ||| |||: |||  
 QY 253 -----TCCCGCAAAAGAGAGAGTGGACGTGACAGATTGGAT 288  
 |||:||||| |||: ||| |||: |||  
 Db 149 LysTyrleuLysLeuAspThrHisSerArgLysArgLysArgLysArgLysArgLys 168  
 |||:||||| |||: ||| |||: |||  
 QY 289 CCATTCCTGTGTGAAGTAATTTGTGACAGATGAACTTCAGTTAAATTAATGT 339  
 |||:||||| |||: ||| |||: |||  
 Db 169 AsnLysCysCysHisValGlyCysThrLysArgSerleuAlaArgPheCys 185  
 |||:||||| |||: ||| |||: |||

## RESULT 6

147053  
 relaxin B,C and A chains - horse (fragment)  
 C:Species: Equus sp.  
 C:Date: 04-Sep-1997 #sequence.revision 07-Nov-1997 #text\_change 16-Jun-1999  
 C:Accession: I47053  
 R:Klonisch, T.; Ryan, P.L.; Yamashiro, S.; Porter, D.G.  
 Biol. Reprod. 52, 1307-1315, 1995  
 A:Title: Partial complementary deoxyribonucleic acid cloning of equine relaxin messenger  
 A:Reference number: I47053; MUID:95359320; PMID:7543293  
 A:Accession: I47053  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-143 <KIO>  
 A:Cross-references: GB:S78800; NID:q1042059; PIDN:AAB35036.1; PID:q1042060  
 C:Superfamily: Insulin

## Alignment Scores:

Pred. No.: 0.00403 Length: 143  
 Score: 103.50 Matches: 30  
 Percent Similarity: 51.76% Conservative: 14  
 Best Local Similarity: 35.29% Mismatches: 34  
 Query Match: 16.30% Indels: 7

DB: 2 Gaps: 3

US-09-518-842-1\_COPY\_76\_417 (1-342) x I47053 (1-143)

QY 7 CTGAGGGATGTGTCTCCCGCATTTGGAAACACTGTGTCATATTGCCCATGCT--- 63  
 |||:||||| |||: ||| |||: |||  
 Db 1 IleLysAlaCysGlyArgGluLeuAlaArgLeuArgLysGlySerLeuSer 20  
 |||:||||| |||: ||| |||: |||  
 QY 64 ---GAGAGACATTC-----ACCACCAACCCAGAGAGGTGGCTGCTGGAATCTGACCG 114  
 |||:||||| |||: ||| |||: |||  
 Db 21 TrpLysThrValLeuArgLeuGluGluProlyl-----LeuGluValGlyGln 37  
 |||:||||| |||: ||| |||: |||  
 QY 115 CCCAAGAAAGTGTGTCAACCTCCACACCAAGATGACACCTTAGGTACACATCA 174  
 |||:||||| |||: ||| |||: |||  
 Db 38 ProvalGluLeuValSerSerSerLysArgLysArgLysArgLysArgLysArgLys 57  
 |||:||||| |||: ||| |||: |||  
 QY 175 GAATTCATCTCTAATTGTGCACAGAGCTGAAGAACCACTGCTGAGGCGACCATCA 234  
 |||:||||| |||: ||| |||: |||  
 Db 58 GlyLeuAsnSerAsnLeuProLysGluGlnLysAlaThrLeuSerGluArgLysProser 77  
 |||:||||| |||: ||| |||: |||  
 QY 235 TTGAGCAATTAATCA 249  
 |||:||||| |||: ||| |||: |||  
 Db 78 TrpArgGluLeu 82  
 |||:||||| |||: ||| |||: |||

## RESULT 7

144559  
 relaxin 1 precursor - human  
 N:Alternate names: preprorelaxin 1  
 C:Species: Homo sapiens (man)  
 C:Date: 03-Feb-1994 #sequence.revision 07-Oct-1994 #text\_change 18-Jun-1999  
 C:Accession: B05092; A44559  
 R:Hudson, P.; John, M.; Crawford, R.; Haralambidis, J.; Scanlon, D.; Gorman, J.; Treg  
 EMBO J. 3, 2333-2339, 1984  
 A:Title: Relaxin gene expression in human ovaries and the predicted structure of a hu  
 A:Reference number: A05092; MUID:85051298; PMID:6548702  
 A:Accession: B05092  
 A:Molecule type: DNA  
 A:Residues: 1-185 <HUI>  
 A:Cross-references: GB:X00949; NID:935932; PIDN:CA25461.1; PID:935933  
 R:Hudson, P.; Haley, J.; John, M.; Cronk, M.; Crawford, R.; Haralambidis, J.; Tregear  
 Nature 301, 628-631, 1983  
 A:Title: Structure of a genomic clone encoding biologically active human relaxin.  
 A:Reference number: A44559; MUID:83141755; PMID:6298628  
 A:Accession: A44559  
 A:Molecule type: DNA  
 A:Residues: 1-185 <HUI>  
 A:Cross-references: GB:X00949; NID:935932; PIDN:CA25461.1; PID:935933  
 C:Comment: Relaxin is an ovarian hormone that acts with estrogen to produce dilation  
 disulfide bonds.  
 C:Genetics:  
 A:Gene: GDB:RLN1  
 A:Cross-references: GDB:119552; OMIM:179730  
 A:Map position: 9pter-9q12  
 C:Superfamily: Insulin  
 C:Keywords: hormone; ovary  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-57/Domain: relaxin 1 chain B #status predicted <BCH>  
 F:58-157/Domain: relaxin 1 chain A #status predicted <ACH>  
 F:162-185/Domain: relaxin 1 connecting C peptide #status predicted <CPEP>  
 F:35-172/47-185/171-176/Disulfide bonds: #status predicted <ACH>  
 F:162-185/Domain: relaxin 1 chain A #status predicted <ACH>

## Alignment Scores:

Pred. No.: 0.00412 Length: 185  
 Score: 103.50 Matches: 44  
 Percent Similarity: 35.67% Conservative: 12  
 Best Local Similarity: 28.03% Mismatches: 44  
 Query Match: 16.30% Indels: 57  
 DB: 1 Gaps: 6

US-09-518-842-1\_COPY\_76\_417 (1-342) x A44559 (1-185)

QY 1 GCAGAGCTGAGGGAGTGTGT---CCCCGATTGGAAACACTGTGTCATATTGCCCC 57



```

Qy      277 -----CAACGATTATCATTCCTGTGAAGAAAT-----
              |||||  |||  :
Db      527 GILuSerThrHisArgGlnHisSerGlyScyGlyIleValSerGlyLysLeuGln 546
              - - - - -TGTGACGATGGAACCTTCATT 330
Qy      310 -----
              |||||  |||||
Db      547 ArgMetIleMetGlnHisSerAspAsnGlyValSerVal 559

RESULT 10
S42783
relaxin 1 precursor - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
C:Accession: S42783
R:Evans, B.B.
submitted to the EMBL Data Library, November 1993
A:Reference number: S42776
A:Accession: S42783
A:Molecule type: mRNA
A:Residues: 1-166 <EYVA>
A:Cross-references: EMBL:Z27225; NID:g415996; PIDN:CAA81739.1; PID:g415997
C:Genetics:
A:Gene: rlx1
C:Superfamily: insulin
C:Keywords: disulfide bond; hormone
F:1-5/Domain: signal sequence (fragment) #status predicted <SIG>
F:6-166/Product: relaxin 1 #status predicted <MAT>

```

Alignment Scores:	
pred. No.:	0.00881
Score:	100.50
Percent Similarity:	36.94%
Best Local Similarity:	26.11%
Query Match:	15.83%
DB: *	2
US-09-518-842-1_COPY_76_417 (1-342) x S42783 (1-166)	
	Length: 166
	Matches: 41
	Conservative: 17
	Mismatches: 42
	Indels: 57
	Gaps: 7

QY	1	GCAGACGTGAGGGGATGTGT---	CCCCATTGTGGAAAACACTTCGTCAATATTTGGCCC	57
Db	23	AlaIInIleIleAlaIleCysGlymetSerThrItpSerLysArgSerLeuSerGlnGluAsp	42	
QY	58	ATGCGCTGAGAGACATTTCACACACCACCCAGAGGGGTGGCTGGAACTTGACGTC	117	
Db	43	AlaProGln-----	ThrPro-----	ArgPro 49
QY	118	---AAGCAATGTGTCTCAACCTTCACAAACAAAGAATGGACAGCCTTAGTACGACATCA	174	
Db	50	ValAlaIleIleValIleProSerPheIleAsnLysAspThrGluThrIleIleIleMetLeu	69	
QY	175	GAATTCATTCCTAATTTGTGCACAGAGCTGAGAGAAACCACTGCTCGAAGCGCCACCATCA	234	
Db	70	GluPheIleAlaAsnLeuProProGlnIleLysAlaIleAlaLeuSerGluArgGlnProSer	89	
QY	235	TTG-----	---	237
Db	90	LeuProGlnProGlnGlnIleValIleProAlaIleLysAspSerAsnLeuSerPheGluGlu	109	
QY	238	---AAGAAATATACTTTCCTCCGCCAATAAAGACAGT---	270	
Db	110	PheLysLysLeuIleArgAsnArgGlnSerGluAlaIleAspSerAsnProSerGluLeu	129	
QY	271	-----GACGCTCACAGATTGATCCATTC-----	294	
Db	130	LysTyrIleuGluLeuAspThrIleIleSerGlnLysLysArgGlnProTyrValAlaLeuPhe	149	
QY	295	-----TGTGTGAAGTAATTTGTGACGATGACGAACTTCAGTTAAATATATGT	339	
Db	150	GluLysCysCysLeuIleGlyCysThrLysArgSerLeuAlaAsnTyrCys	166	

```

relaxin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 24-Sep-1999
C:Accession: A01614
R:Hudson, P.; Haley, J.; Cronk, M.; Shine, J.; Niall, H.
NATURE 291, 127-131, 1981
A>Title: Molecular cloning and characterization of cDNA sequences coding for rat relaxin
A:Reference number: A01614; MUID:81197624; PMID:7231533
A:Accession: A01614
A:Molecule type: mRNA.
A:Residues: 1-186 <HDU>
A:Cross-references: GB:J00780; GB:M25468; NID:g206606; PIDN:AAAA2029.1; PID:g206607
C:Superfamily: Insulin
C:Keywords: hormone; ovary; pyroglutamic acid
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-186/Product: relaxin chain B #status predicted <RXB>
F:123-57/Domain: relaxin chain B #status predicted <RM>
F:123-57/163-186/Product: relaxin #status predicted <RM>
F:58-162/Domain: relaxin connecting C peptide #status predicted <RCX>
F:163-186/Domain: relaxin chain A #status predicted <RA>
F:36-173,48-186,172-177/Disulfide bonds: #status predicted
F:163/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted

```

Alignment Scores:	
Pred. No.:	Length:
0.0192	186
Score:	Matches: 38
Percent Similarity:	Conservative: 31.79%
Best Local Similarity:	Mismatches: 48
Query Match:	Indels: 55
DB:	Gaps: 1
	4

US-09-518-842-1\_COPY\_76\_417 (1-342) x RXRT (1-186)

```

QY      16 TGTGGTCCCGCATTTGGCAAAACACTTGCTGTCATAT---TGCCCCAGTCCGTGAAGAACA 72
       |||||          ||:::         |||||          |||   |||
Db      48 CysgIyAlaserValGIyArgLeuAlalLeuSerIngluProLaPro-----64
QY      73 TTCACCACCCAGGAGGGGTGCCTGTGGAATTCTGGACGCTCCCAAAGAAATGCTCA 132
       |||:::|||||          |||
Db      65 -----LeuAlaraglInlarThrAlagluValaPro 75
QY      133 ACCTTCAACACAAAGATGGACAGCCTTAGAGTACGACATCACATTCATTCTCAATTGG 192
       ::::| | | | | | | : : : | | | : : : | | | | | | | |
Db      76 SerPheIIeAnLysAspAlaGluprPheaspMethrLeuLysCysLeuProAsnLeu 95
QY      193 TCACCAAGAGCTGGAACCAACACACTGTCTGCAAGCGCA-----228
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      96 SerGIuglUArGLVSAIAlaLeuSerCIuIdlyArgAlarProrPheproGIuleGlnGln 115
QY      228 -----228
Db      116 HisAlaProAlaleuSerAspSerValValSerLeuGIuGIyPheLysLysThrPheHis 135
QY      229 -----CCATCATTTGAGAAATAA-----246
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      136 AsnGIuLeuGIyGluAlaGluAspCIyLyprorProGIuleLysTYrLeuGIySerAsp 155
QY      247 ATACTTTCGCCGCAAAAAGAGAGTGGACGTCACAGATTGTATCCATTCTGTGTGAAGTA 306
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      156 AlaGlnSerArgLyLysArgGlnSerGIyleuAlaleuLeuSerGIuGlnCysHISile 175
QY      307 ATTGTGACGATGACACTTCAGTTAAATTATGT 339
       ||| : : : | | | | | | | |
Db      176 GLyCysThrArgArgSerIleAlaLysLysCys 186

RESULT 12
E83259
hypothetical protein PA3089 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Spectes: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Acession: E83259
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; L
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Tarbidy, K.; L
.: Lory, S.; Olson, M.V.
```

R.; Lee, Y. A.; Bryant-Greenwood, G. D.; Mandel, M.; Greenwood, F. C.

Endocrinology 130, 1165-1172, 1992

A:Title: The complementary deoxyribonucleic acid sequence of guinea pig endometrial pro

A:Reference number: A49194; MID:92164503; PMID:1537282

A:Accession: A49194

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-160 &lt;LEE&gt;

A:Cross-references: GB:S85964; NID:9246391; PIDN:AAB21586.1; PID:9246392

C:Superfamily: Insulin

## Alignment Scores:

Pred. No.:	4,14	Length:	160
Score:	76.50	Matches:	31
Percent Similarity:	35.09%	Conservative:	9
Best Local Similarity:	27.19%	Mismatches:	31
Query Match:	12.05%	Indels:	43
DB:	2	Gaps:	3

US-09-518-842-1\_COPY\_76\_417 (1-342) x A49194 (1-160)

```
QY 97 CTGCTGGAATCTGACGTCCTCCAAAGAAATGCTCAACCTCCACACACAAAGATGCACAA 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 ILeuclYserGlyInserAlaGluIleMetProSerSerIleAsnLysGluValAsp 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 GCCTTAGTACGACATCAGAAATTCCTTAATTTGGACACAGAGCTGAAGAAACCACTG 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 SerLeuAsnMetLeu---GluSerIleAlaAsnLeuProGluGluLeuArgAlaMetLeu 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 TCTGAAGCGCAGCCATCA----- 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 ProGluLysGlnProSerSerProGlnLeuGlnGlnIleTyValProAlaLeuLysAsnSer 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 235 -----TTGAGAAATAATTAATCTTCCCGC----- 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 AsnValAlaValLysGluLeuAsnLysIleIleArgGlyArgGlnGluAlaGluAsp 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 259 -----AAAAAGAGAGT 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 AsnSerHisSerLeuLysAspPheAsnLeuAsnIleTySerProLysLysArgGln 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 GGACGTCACAGATTGATTCATTCTGTGTGAAGTAATTGT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 LeuAspMetThrValSerGluLysCysCysGlnValGlyCys 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 14, 2003, 19:46:49  
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: June 14, 2003, 19:35:45 ; Search time 35 Seconds

(without alignments)  
4026.749 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 635

Sequence: 1 GCAGAGCTGAGGAGATGTGG.....CTTCAGTAAATATATCA 342

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool/US09518842/runat\_14062003\_175811.12270/app.query.fasta.1.519  
-DB=SPTRMBL.21 -QEMT=fastan -SUFFIX=n2p.ispt -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=Blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09518842@cgn.1.1.138 -runat\_14062003\_175811.12270 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL.21:  
1: sp.archaea:  
2: sp.bacteria:  
3: sp.fungi:  
4: sp.human:  
5: sp.invertebrate:  
6: sp.mammal:  
7: sp.mhc:  
8: sp.organelle:  
9: sp.phage:  
10: sp.plant:  
11: sp.protozoa:  
12: sp.virus:  
13: sp.vertibrate:  
14: sp.unclassified:  
15: sp.virus:  
16: sp.bacteriap:  
17: sp.archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	113.5	17.9	199 6 Q9N0T9	Q9N0T9 camelus dro

2	104	15.4	180	6	Q9MYK8	Q9MYK8 felis silve
3	101	15.9	560	5	Q22328	Q22328 caenorhabdi
4	83	13.1	321	16	Q9H2B9	Q9H2B9 pseudomonas
5	82	12.9	795	11	Q9DBY0	Q9DBY0 mus muscun
6	80.5	12.7	889	10	Q9FHH9	Q9FHH9 aradidopsi
7	78	13.0	754	16	Q8ZOC3	Q8ZOC3 salmonella
8	78	13.0	754	16	Q8ZOC3	Q8ZOC3 salmonella
9	77	12.1	457	16	Q8ZB58	Q8ZB58 yersinia pe
10	75.5	12.6	718	6	Q9BE45	Q9BE45 bos taurus
11	75.5	12.6	793	4	Q9H522	Q9H522 homo sapien
12	74.5	11.7	752	10	Q94H57	Q94H57 oryza sativ
13	74	11.7	475	10	Q9FTX0	Q9FTX0 oryza sativ
14	74	11.7	502	10	Q942M3	Q942M3 oryza sativ
15	73.5	11.6	273	16	Q95X55	Q95X55 staphylococ
16	73.5	12.2	402	10	Q82202	Q82202 aradidopsi
17	73	11.5	590	4	Q96G63	Q96G63 homo sapien
18	73	11.5	636	4	Q9H9B6	Q9H9B6 homo sapien
19	73	11.5	695	4	Q9H062	Q9H062 homo sapien
20	73	11.5	747	4	Q96A55	Q96A55 homo sapien
21	73	11.5	747	4	Q96A55	Q96A55 homo sapien
22	73	11.5	747	4	Q96A55	Q96A55 homo sapien
23	72.5	11.4	1435	3	Q03291	Q03291 saccharomyc
24	72.5	11.4	1432	5	Q95ZV4	Q95ZV4 caenorhabdi
25	72	11.3	391	2	Q44032	Q44032 caenorhabdi
26	72	11.3	483	16	Q8YBNO	Q8YBNO ralslonia m
27	72	12.0	506	17	Q96X64	Q96X64 bruceella me
28	72	11.3	752	3	Q9UR72	Q9UR72 schizosacch
29	71.5	11.9	1217	5	P91457	P91457 caenorhabdi
30	71.5	11.3	1530	2	Q930E8	Q930E8 chlamydia t
31	71	11.2	204	4	Q9BRP8	Q9BRP8 chlamydia t
32	71	11.8	206	3	Q99395	Q99395 saccharomyc
33	71	11.2	327	5	Q9W028	Q9W028 drosophila
34	71	11.2	697	5	Q9V011	Q9V011 drosophila
35	71	11.2	836	12	Q66400	Q66400 duck hepati
36	71	11.2	1417	12	Q89241	Q89241 western equ
37	71	11.2	1490	12	Q88791	Q88791 western equ
38	70.5	11.1	37	3	P87229	P87229 schizosacch
39	70.5	11.1	476	5	Q961W9	Q961W9 drosophila
40	70.5	11.8	766	10	Q9LNC4	Q9LNC4 aradidopsi
41	70.5	11.1	1037	3	Q8WZK1	Q8WZK1 neurospora
42	70.5	11.8	2015	5	Q9VAS6	Q9VAS6 drosophila
43	70	11.0	542	5	Q9NJD9	Q9NJD9 onchocerca
44	70	11.7	693	4	Q9N82	Q9N82 homo sapien
45	70	11.7	747	4	Q9Y6A1	Q9Y6A1 homo sapien

## ALIGNMENTS

RESULT 1

Q9N0T9 PRELIMINARY: PRT: 199 AA.

AC Q9N0T9; 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Preprorelaxin.  
OS Camelus dromedarius (Dromedary) (Arabian camel).  
OC Eukaryota; Metazoa; Chordata; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
OX NCBI\_TaxID=9838;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20191540; PubMed=10727251;  
RA Hombach-Klonisch S., Abd-Elmeleu M., Skidmore J.A., Leiser R.,  
RT "Ruminant relaxin in the pregnant one-humped camel."  
RL Biol. Reprod. 62:839-846(2000).  
CC 1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
DR EMBL: AF254739; AAF6741.1; -  
DR HSSP: P01346; IRLX.  
DR InterPro: IPR004825; Ins/IGF/relax.  
DR Pfam: PF00049; Insulin; I.

DR SMART: SM00078; IIGF: 1.  
DR PROSITE: PS00262; INSULIN: 1.  
SQ SEQUENCE 199 AA; 22386 MW; F69BBD7E4F4A8089 CRC64;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000149	199	42	10	59	4
Percent Similarity:	113.50					
Best Local Similarity:	30.95%					
Query Match:	25.00%					
	17.87%					

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q9N079 (1-199)

```

QY 7 CTGAGGGGATGTGGTCCCGATTGGAAAAACCTTGTCATATTGC----- 54
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 VallysalacysgllyargluleuValargleuTrpIleGlullecysglserValser 51
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 55 -----CCCATGGCTGAGAGACATTCACCCACCCGAGGAGGCGTG----- 99
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 TrpGlyArgProAlaProArgProAlaProArgProAlaProLysProAlaLeuArgPro 71
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 100 -----CTGAATCTGACGTCACCAAGAAATGAGTGA 132
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 AlaLeuSerGlnAspLysLysProArgLeuArgSerGlyProProAlaGluLeuMetPro 91
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 ACCTCCACACAAAGATGAGACAGCCATTAGTACACATCAGATTCATTCTAATTGG 192
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 SerSerLeuThrLysAspAlaGluThrLeuThrMetLeuGluPheThrProAsnLeu 111
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 TCACACAGAGCTGAAGAAACCATGTCGAGGAGGCACCATCA----- 234
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 ProGlnGluLeuThrAlaThrLeuSerGlnArgGlnProSerAlaGluProGlnGlnPro 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 -----TGAGAGAAATAATTAATCTTCCCGC--- 258
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 132 AlaLeuLysAspSerAsnLeuAsnPheGluGluPheLysIleIlePheAspArgGln 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 ----- 258
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 AsnGluGluLysAspGluSerLeuSerGluLeuLysAsnLeuGlyLeuAspLysHisSer 171
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 -----AAAAAGAGAGTGGACGCTCACAGATTGATCCATTCTGTGAGAGTAATTGTGAC 315
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 GluLysLysArgGlnLeuGlnMetThrLeuGlyGluArgCysGlnLysGlyCysSer 191
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 GATGGAACTTCAGTTAATTATGCT 339
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 ArgLysGluMetAlaThrAlaCys 199
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 2

Q9MYK8 PRELIMINARY: PRT: 180 AA.

AC Q9MYK8;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Relaxin.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCB1\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99115216; PubMed=9915995;  
RA Klonisch T., Hombach-Klonisch S., Froehlich C., Kaufold J.,  
RA Steger K., Huppertz B., Fischer B.,  
RA "Nucleic acid sequence of feline preprorelaxin and its localization  
RT within the feline placenta."  
RU Biol. Reprod. 60:305-311(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Hombach-Klonisch S., Klonisch T.;

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IIGF/RELAXIN FAMILY.

DR EMBL: AF233688; AAF60303.1; -  
DR HSSP: P04090; 6RLX.  
DR InterPro: IPR004825; Ins/IIGF/relax.  
DR Pfam: PF00049; Insulin.1.  
DR SMART: SM00078; IIGF: 1.  
DR PROSITE: PS00262; INSULIN: 1.  
SQ SEQUENCE 180 AA; 20360 MW; 4C2CF371C698A9F9 CRC64;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.00194	180	40	13	47	62
Percent Similarity:	104.00					
Best Local Similarity:	32.72%					
Query Match:	24.69%					
	16.38%					

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q9MYK8 (1-180)

```

QY 7 CTGAGGGGATGTGGTCCCGATTGGAAACACTTGTCATATTGCCCATG----- 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 LeuLysAlaCysGlyArgGluPheValArgLeuGlnIleArgIleCysGlySerLeuSer 49
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 -----CCTGAAGAACATTCACACCACCCAGCA 90
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 50 TrpGlyLysSerSerGlnGlnHisArgGluProArgGlnAlaProAlaLeuPro--- 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 91 GGGTGGCTGCTGGAATCTGGACGTCACCAAGAAATGTGTACCTCCACCAACAAAGAT 150
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 -----GluLeuValSerSerLeuThrSerGly 78
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 151 GGACAGGCTTAGTACGACACATCAGAAATTCATTCTAATTGTACACAGAGCTGAAGAA 210
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 79 AlaGluAlaLeuAsnGlyMetLeuGlnIuTrIleProAspLeuProGlnGluLeuLysAla 98
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 CCACGTGTGAAGGGACGCATCA----- 234
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 ThrLeuSerGluArgGluProSerPheArgGluLeuGlnProSerLeuLysAspSerAsn 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 -----TGAGAGAAATAATTAATCTT----- 252
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 LeuAsnLeuGluGluValGluLysSerIleLeuGlyArgGlnAsnGluAlaGluAspGln 138
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 -----TCCGCCAAAAGAGAGTGCA 273
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 SerLeuSerGlnLeuGlyArgSerArgLeuAspAlaHisSerArgIleLysArgSerAsp 158
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 274 CGTCACAGATTGATTCATTCTGTTGTGAAGTAATTGTGACGATGCACTTCAGTTAA 333
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 TyrIleArgTyrSerAspArgCysAsnValGlyCysThrArgLysGluLeuAlaAsp 178
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 334 TTAATG 339
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 LeuCys 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 3

Q22328 PRELIMINARY: PRT: 560 AA.

AC Q22328;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypochemical 61.6 kDa protein.  
GN T07H6.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCB1\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;

Alignment Scores:	
Pred. No.:	0.00564
Score:	101.00
Percent Similarity:	38.56%
Best Local Similarity:	28.10%
Query Match:	15.91%
DB:	5
Gaps:	5
Length:	560
Matches:	43
Conservative:	16
Mismatches:	38
Indels:	56

QY	19	GGTCCCGGATTGGAAACACCTTCGTCGATATTTGGCCCATCCCTGGAACACATTACAC	78
	1414		
Db	414	GlyProPromet---LysCysLeuProSerTyrCysGluHisProSerLysThrTyrLys	433
QY	79	ACACACCCGAGGAGGTCGTCGTCG-----	102
	433		
Db	433	ThrLeuProGlyGlnIleLeuLeuGlnGlyIleLeuGlyAlaTyrGluPheGlnSer	455
QY	103	-----GAATCTGGAGCTCCCAAGAATGCTGTCACTCCACAACAAA	147
	453		
Db	453	TyrIleGlnLysValGlnGluGlnGlyTyr-----AlaIleSerPheGlnCysGlyLys	465
QY	148	GATGACACAGCCTTAGGT-----	189
	470		
Db	470	GlyAsnTyrLeuIleGlyProProlLysAlaThrCysValAsnGlyGluTyrMetProlLys	489
QY	190	TTTCACACACAGGCTGMAAACACACACCTGTGTGAAGGACCATTCATTGAG--AAATA	246
	490		
Db	490	ValSerPro-----LysCysValSerGlnThrHisPrometIleGlnGlyLysIle	506
QY	247	ATACTTCCCGCAAAAAGAGAGT-----GGAGCT	276
	507		
Db	507	LeuTyrPaspArgLysLysArgSerLeuProGlnAlaValAlaArgLysTyrValAspAsp	526
QY	277	-----CACGATTTCATTCATCTGTTCGTAAGTAAT-----	309
	527		
Db	527	GluLeuSerThrHisArgGlnHisSerGlyLysCysGlyIleValSerGlyLysLeuGlu	546
QY	310	-----TGTGACGATGGACTTCAGTT	330
	547		
Db	547	ArgMetIleMetGlnHisSerAspAsnGlyAlaSerVal	559
RESULT 4			
Q9H2B9			
ID	Q9H2B9	PRELIMINARY:	PRT: 321 AA.
AC	Q9H2B9;		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		

DT 01-OCT-2001 (TRMBLrel. 18, Last annotation update)  
 GN Hypothetical protein PA3089.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10964043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garner R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE0047733; AAC60477.1; "-"  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 321 AA; 36640 MW; 3EBB5FDD3DBF268F CRC64;  
  
 Alignment Scores:  
 Pred. No.: 0.671 Length: 321  
 Score: 83.00 Matches: 17  
 Percent Similarity: 61.54% Conservative: 17  
 Best Local Similarity: 43.59% Mismatches: 6  
 Query Match: 13.07% Indels: 9  
 Gaps: 2  
  
 US-09-518-842-1\_COPY\_76\_417 (1-342) x Q9H2B9 (1-321)  
 QY 19 GGTCGCCGCGATTGGAAACACCTGCGTCAT-----ATTGCCCATG 60  
 Db 234 GLYPROGlnhrhrlgylglnhistrpleu-HisIeuSerArgGlnHistrpleuAlaProI 253  
 QY 61 CCTGAGAAGACATTCACCA-----CCACCCCGAGAGGAGTGCTCGTAATTC 109  
 Db 253 aetlvtgrgArgltnasprleuaspProleuaspProIargrgAlacysrtpsanlle 271  
  
 RESULT 5  
 Q9DBYO PRELIMINARY; PRT; 795 AA.  
 AC Q9DBYO;  
 DT 01-JUN-2001 (TRMBLrel. 17, Created)  
 DT 01-JUN-2001 (TRMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TRMBLrel. 21, Last annotation update)  
 DE 1200010K03RIK protein.  
 GN 1200010K03RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LUNG;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Alawra K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Stabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher N., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,





RA Prentice M. B., Sepalia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarriga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltham T., Hamlin N., Holroyd S., Jagels K., Katlyshev A.V.,  
RA Leather S., Moutle S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrill B.G.,  
RA "Genome sequence of Yersinia pestis, the causative agent of plague",  
RT Nature 413:523-527(2001).  
RL EMBL; A0414157; CAC92795.1; -.  
DR MEROPS; S01.274; -.  
DR InterPro; IPR001478; PDZ.  
DR InterPro; IPR001940; Protease2C.  
DR InterPro; IPR001254; Ser\_\_protease\_Try.  
DR Pfam; PF00595; PDZ; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00834; PROTEASES2C.  
DR SMART; SM00228; PDZ; 2.  
DR PROSITE; PS50106; PDZ; 2.  
KW Hydroxylase, Complete proteome.  
SO SEQUENCE 457 AA; 47399 MW; B89E79B8229FC86F CMC64;

Alignment Scores:	
Pred. No.:	3.72
Score:	77.00
Percent Similarity:	47.89%
Best Local Similarity:	35.21%
Query Matchn:	12.13%
DB:	16
Length:	45
Matches:	25
Conservative:	9
Mismatches:	27
Indels:	10
Gaps:	3

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q82B58 (1-457)

[illegible]

RESULT 10

```

ID      Q9BE45; PRELIMINARY; PRT; 7A AA.
AC      Q9BE45;
DT      01-JUN-2001 (TREMBlrel, 17, Created)
DT      01-JUN-2001 (TREMBlrel, 17, Last sequence update)
DT      01-DEC-2001 (TREMBlrel, 19, Last annotation update)
DE      MAIL.
GN      MAIL.
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea
OC      Bovidae; Bovinae; Bos.
OX      NCBI_TaxID=9913;
RX      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=HOLSTEIN; TISSUE=BLOOD;
RA      Yamaji D., Kitamura H., Morimatsu M., Shina T., Kanehira K.,
RA      Fujikura D., Saito M.,
RT      "Bos taurus mRNA for MAIL, complete cds."
RL      Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR      HMBB; AB058410; BAB39767.1; -.
DR      HSSB; P25963; INFEI.
DR      InterPro: IPR002110; ANK.
DR      Pfam: PF00023; ank_6
DR      PRINTS: PRO1415; ANKTHIN.

```

```
DR SMART; SMO0248; ANK: 6
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 718 AA; 78247 MW; CB8D8F6A994A9E00 CRC64;
```

Alignment Scores:	
Pred. No.:	6, 18
Score:	75, 50
Percent Similarity:	39, 05%
Best Local Similarity:	28, 57%
Query Match:	12, 58%
DB:	6
	Gaps:
	5
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q9BE45 (1-718)

```

QY      272  CCACCTCTCTTTTGGGGGAAAGATATATTTCTTCATGATAGGCTGGCCCTGACGAGT 213
      ||| ||||| |||||
Db      324  ProHisIleuPheGlyArgIlePro-----GlnPheCysPro--AspGln 337
      ||| ||||| |||||
QY      212  GGTTC-----TTCAGCTCTGGTGACAAATTAGGAATTCGATGCTGACTAAG 159
      ||| ||| |||
Db      338  SerPheIleProIleuLeuSerIleProAlaGlnSerGlnAsnIleAlaValIleProGln 357
      ||| ||| |||
QY      158  GCTGTGCATCTTTGGTGTGGAGCTTGACACC----- 126
      ||||| ||||| |||||
Db      358  ThrIleProSerValGlnGlnGlnIleAspThrIleHisLeuGlnAsnPheSerLeuMetPro 377
      ||||| ||||| |||||
QY      125  -----ATTCTTTGGAGCTCCAGATTCCAGACGCCACCTCTCGGGGTGGTG 78
      ||||| ||||| ||||| |||||
Db      378  ProAspThrCysGlnAlaLeuAlaIleArgProAspAlaSerSerThrProLeuSerThrPro 397
      ||||| ||||| ||||| |||||
QY      77  GTGATGTCTTTCACAGCATGGGCGCATATGACAGCAAGTGTTCACCAATCGGGAGCA 18
      |||
Db      398  LeuPro-----PheProAsnLeuGlyGly 405
      ||||| |||||
QY      17  CATGCCCTCAGCTCT 3
      ||||| |||||
Db      406  AsnProMetSerIle 410

```

## RESULT 11

```

ID      09H52Z      PRELIMINARY;      PRT;      793 AA.
AC      Q9H52Z;
DT      01-MAR-2001 (TREMBlrel. 16, Created)
DT      01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE      CDNA: FLJ22757 fis, clone KAIAB083.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=ILEAL MUCOSA;
RA      Kakekimi T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura T.,
RA      Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ota T., Suzuki
RA      Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y., Ota T., Suzuki
RA      Obaashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA      Isogai T., Sugano S.,
RT      "NEO human cDNA sequencing project."
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AK026410; BAB5475.1; -.
DR      InterPro: IPR005112; dENN.
DR      InterPro: IPR001194; dENN.
DR      InterPro: IPR005113; dENN.
DR      Pfam: PF03455; dENN; 1.
DR      Pfam: PF02141; dENN; 1.
DR      Pfam: PF03456; dENN; 1.
SQ      SEQUENCE 793 AA; 86230 MW; 3E37ACA836F24AF4 CMC64;

```

Alignment Scores: 6.32 Length: 7932  
Pred. No.:

Score: 75.50  
 Percent Similarity: 42.11%  
 Best Local Similarity: 27.37%  
 Query Match: 12.58%  
 DB: 4  
 Gaps: 4

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q9H572 (1-793)

QY 236 AATGATGCTGCTC-----TCAGACAGTGTCTTCTGAGTCTGCTGAC 192  
 DB 536 ASPLUGLYCSPROTRIPALAGLUALALenAspSerPheLeuGlySerLysGlu 555  
 QY 191 AATATGAGATGATTCATCTGATCTGACTTAAGCTGTGCATCTTGTGTGAGAGTT 132  
 DB 556 GIuleAspLeuSerGluLeu-----Leu 564  
 QY 131 GACACATTTCTTGTGGAGCTCAGATTCACAGCAGCAGCTCTCT----- 87  
 DB 565 ASpSerLeuSerMetGlyAlaLysSerAlaGlySerLeuArgProSerGlnSerLeuAsp 584  
 QY 86 -----GGGCTGTGTGTGATGCTCTTCAGAGCAGCAGG-----CAA 51  
 DB 585 CysCysHisArgGlyAspLeuAspSerCysPheSerLeuProAsnLleLeuArgTrrpIn 604  
 QY 50 TATGACAGCAATGTTTCCAAATCGGGACACATCCCTCAGC 6  
 DB 605 ProAspAspLysLysLeuProGluProGluProGluProLysLeuSer 619

## RESULT 12

Q94H57 PRELIMINARY: PRT: 752 AA.  
 AC 094H57:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DE Putative retrotransposon protein.  
 OS Oryza sativa (Rice)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriaristolidae; Oryzae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
 RA Brenner M., Burgess S., Hance M., Shwartsbeyn M., Tsitlin T.,  
 RA Riggins F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,  
 RA Uterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,  
 RA White O., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSNBNB0057P11 genomic sequence."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AC084767; AAK72287.1;  
 SQ SEQUENCE 752 AA; 83920 MW; 88AC2AEC1157DB98 CRC64;

## Alignment Scores:

Score: 8.2  
 Percent Similarity: 74.50  
 Best Local Similarity: 48.61%  
 Query Match: 30.56%  
 DB: 10  
 Gaps: 2

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q94H57 (1-752)

QY 47 CATATGCCCCATCCCTG-----AGAAGACATTCACACACCCAGAGGTGGCTGC 100  
 DB 280 HisLeuLeuProValValProTyrAsnAlaPheAsnProProProGluLupheserLeu 299  
 QY 101 TGGATCTGGACGCTCCCAAGAAATGCTCAACCTCCAC-AACAAGATGACAAGCC 159  
 DB 300 MetLysSerAspProProThnAlaGlnCysArgSerProArgArgGlnThrGlyGlnAla 319  
 QY 160 TTAGGTACGACATCAGATTCATCTCAATTGTGCACACAGAGCTGAAGAACAACGTCT 219

DB 320 SerGly---SerProLysPheGlnProThnValGlnProSerSerProLysProThnAsn 338  
 QY 220 GAAGGACAGCCATCATTCAGAAATAATACTTCC 255  
 DB 339 GlnSerGlySerGlnLysArgLysLeuValLeuSer 350

## RESULT 13

Q9FTX0 PRELIMINARY: PRT: 475 AA.  
 AC 09FTX0:  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE P0407B12.10 protein.  
 CN P0407B12.10.  
 OS Oryza sativa (Rice)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriaristolidae; Oryzae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0407B12."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AP002843; BAB17173.1;  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 SQ SEQUENCE 475 AA; 53631 MW; A9EA462BF6AB21B3 CRC64;

## Alignment Scores:

Pred. No.: 8.49  
 Score: 74.00  
 Percent Similarity: 44.34%  
 Best Local Similarity: 26.42%  
 Query Match: 11.65%  
 DB: 10  
 Gaps: 3

US-09-518-842-1\_COPY\_76\_417 (1-342) x Q9FTX0 (1-475)

QY 1 GCAGAGCTGAGGCGATGTGTCCTCCGATTTGGAAACACTTGTGTCATATGCCCCATG 60  
 DB 269 SerGlnLeuArgAsnArgAlaLeuGlnPheValLysHisLeuGlyLysThrLeuProLeu 288  
 QY 61 CTGAGAGACATTCACACC-----ACCCAGAGAGG 93  
 DB 289 ThrSerSerAsnLeuSerThrHisHisLysLeuLysLysLeuValValProAlaPro 308  
 QY 94 TGCTCTCGTAATCTCGA-----CGTCCCAAGAAATGCTCAACCTCCAGAG 141  
 DB 309 ArgMetLysGlnLeuAsnAsnGlnGluPheArgProGluAspLeuLeuLeuAlaAsp 328  
 QY 142 AACAAAGATGACACACCTTAGGTACGACATCAGATTCATCTCAATTGTGCACAGAG 201  
 DB 329 SerValAspLysPheSerLeu---MetThrTyrAspPheSerGlyProGlnSerProGly 347  
 QY 202 CTGAACAAACACTGCTGTAAGGCGACGACATCATTCAGAAATAATACTTCCGCAAA 261  
 DB 348 ProSerAlaProLeuSerTrrpLieglnTyrSerLeuLysThrLeuLeuAlaLysGly 367  
 QY 262 AACAGAGTGGACGCTCAG 279  
 DB 368 SerAlaSerHisGlyHis 373

## RESULT 14

Q942M3 PRELIMINARY: PRT: 502 AA.  
 AC 0942M3:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
 DE B1040D09.22 protein.  
 GN B1040D09.22.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC  
 clone:B1040D09.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003328; BAB67973.1; -  
 DR InterPro: IPR001223; Glyco\_hydro.18.  
 DR Pfam: PF00704; Glyco\_hydro.18; 1.  
 SO SEQUENCE 502 AA; 56877 MW; 0B6EA76DF4555660 CRC64;

Alignment Scores:  
 Pred. No.: 8.59 Length: 502  
 Score: 74.00 Matches: 28  
 Percent Similarity: 44.34% Conservative: 19  
 Best Local Similarity: 26.42% Mismatches: 45  
 Query Match: 11.65% Indels: 14  
 DB: 10 Gaps: 3

US-09-518-842-1\_COPY\_76\_417 (1-342) x 0942M3 (1-502)

QY 1 GCAGACCTGAGGAGATGTCGCCGATTTGGAACACTGTCGATATGCCCCCATG 60  
 DB 296 SerGluLeuTrpGlnArgAlaLeuGlnPheValHisLeuGlyThrLeuProLeu 315  
 QY 61 CCGAGAGACATTCACACACC-----ACCCAGAGGCG 93  
 DB 316 ThrSerSerAlaLeuSerThrHisHisLeuLysLeuLeuValValProAlaPro 335  
 QY 94 TGGCTGCTGGAATCTGGA-----CGTCCCAAGAAATGGTGCACACCTCCAC 141  
 DB 336 ArgMetLysGluLeuAsnAsnGlnGluPheArgProGluAspLeuLeuGlnLeuAlaAsp 355  
 QY 142 AAGAAAGACGACACCTTAGGTACGACATCAGAAATTCATTCCTAATTTGTCACCGAG 201  
 DB 356 SerValAspLysPheSerLeu---MetThrTyraSphSerGlyProGlnAsnProGly 374  
 QY 202 CTGAGAGAACCACTGTCGAGGCGCAGCCATCTTAAGAAATAATACTTCCGCCCAA 261  
 DB 375 ProSerAlaProLeuSerTrpIleGlnTyrSerLeuLysThrLeuAlaAlaLysGly 394  
 QY 262 AAGAGAGTGGAGCTCAC 279  
 DB 395 SerAlaSerHisGlyHis 400

RESULT 15

099SX5 PRELIMINARY; PRT; 273 AA.

AC 099SX5;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
 DE NAD synthetase, prefers NH3 over glutamine.  
 GN MADE OR SAV1912 OR SAV1728.  
 OS Staphylococcus aureus (strain Mu50 - ATCC 700699), and  
 OS Staphylococcus aureus (strain N315)  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 aureus.";  
 RL Lancet 357:1225-1240(2001).  
 DR EMBL; AP003363; BAB58074.1; -  
 DR EMBL; AP003135; BAB42998.1; -  
 DR HSSP; P08164; 2MSY.  
 DR InterPro: IPR003694; NAD\_synthase.  
 DR Pfam: PF02540; NAD\_synthase; 1.  
 DR TIGRFAMs: TIGR00552; nade; 1.  
 KW Complete proteome.  
 SO SEQUENCE 273 AA; 30683 MW; EB40D2F38247F08 CRC64;

Alignment Scores:  
 Pred. No.: 8.62 Length: 273  
 Score: 73.50 Matches: 30  
 Percent Similarity: 44.23% Conservative: 16  
 Best Local Similarity: 28.85% Mismatches: 30  
 Query Match: 11.57% Indels: 28  
 DB: 16 Gaps: 5

US-09-518-842-1\_COPY\_76\_417 (1-342) x 099SX5 (1-273)

QY 1 GCAGACCTGAGGAGATGTCGCCGATTTGGAACACTGTCGATATGCCCCCATG 60  
 DB 181 AlaProIlePheGlyLeuAsnLysArgGlnGlyArgGlnLeuLeuAlaTyrLeuGlyAla 200  
 QY 61 CCGAGAGACATTCACACACCCCGAGAGGCGTGGCTGGAATCTGGACGCTCCCAA 120  
 DB 201 ProLysGluLeuTyrGlnLysThrProThraL---AspLeuGlnAspAspLysProGln 219  
 QY 121 GAATGTGTGACCTCCACACACAAAGATGGACACCTTAGTACGACATCAGAAATC 180  
 DB 220 -----LeuProAspLysAlaLeuGlyValThrTyrGluAla 232  
 QY 181 ATTCTCTAT-----TTGTCACACGAG-----CTGAGAGAA 210  
 DB 233 IleAspAsnTyrLeuGlnGlyLysProValThrProGluGlnGlnValIleGlu-As 252  
 QY 211 CCACGTGTCGAAGGCGACCCATCATTTGAAGAAATAAATACTTCCGCCCAAAAGAGAGT 270  
 DB 252 nHisTyrIleArgAsnAlaHis-----LysArgGluLe 263  
 QY 271 GGACGTCACA 280  
 DB 263 uAlaTyrThr 266

Search completed: June 14, 2003, 19:46:07  
 Job time : 41 secs